

**Validierung eines makroskopischen und eines kamerabasierten  
Bewertungssystems der Foot-pad Dermatitis bei Masthühnern  
durch histopathologische Untersuchungen**

von André Piller

Inaugural-Dissertation zur Erlangung der Doktorwürde der Tierärztlichen Fakultät der  
Ludwig-Maximilians-Universität München

**Validierung eines makroskopischen und eines kamerabasierten  
Bewertungssystems der Foot-pad Dermatitis bei Masthühnern  
durch histopathologische Untersuchungen**

von André Piller  
aus Viechtach

München 2020



Aus dem Veterinärwissenschaftlichen Department  
der Tierärztlichen Fakultät  
der Ludwig-Maximilians-Universität München

Lehrstuhl für Tierschutz, Verhaltenskunde, Tierhygiene und Tierhaltung

Arbeit angefertigt unter der Leitung von: Univ.-Prof. Dr. Michael H. Erhard

Mitbetreuung durch: Univ.-Prof. Dr. Helen P. Louton



**Gedruckt mit der Genehmigung der Tierärztlichen Fakultät  
der Ludwig-Maximilians-Universität München**

**Dekan:** Univ.-Prof. Dr. Reinhard K. Straubinger, Ph.D.

**Berichterstatter:** Univ.-Prof. Dr. Michael H. Erhard

**Korreferent:** Univ.-Prof. Dr. Rüdiger Korb

**Tag der Promotion:** 25. Juli 2020



**Für meinen Bruder Patrick  
und  
die zukünftigen Generationen von Masthühnern**





<b>I.</b>	<b>EINLEITUNG .....</b>	<b>9</b>
<b>II.</b>	<b>ERWEITERTE LITERATURÜBERSICHT .....</b>	<b>11</b>
1.	Pathologie und Pathogenese der Foot-pad Dermatitis.....	11
2.	Makroskopische Scoringsysteme für Foot-pad Dermatitis.....	12
3.	Histologische Scoringsysteme für Foot-pad Dermatitis .....	17
4.	Gegenwärtige Erfassung von Foot-pad Dermatitis in Deutschland .....	19
<b>III.</b>	<b>TIERE, MATERIAL UND METHODEN .....</b>	<b>21</b>
1.	Beobachterabgleich.....	22
2.	Datenerhebung am Schlachthof.....	23
2.1.	Verwendete Tiere .....	23
2.2.	Makroskopische und kamerabasierte Beurteilung.....	23
3.	Histologische Beurteilung .....	25
4.	Stichprobenberechnung.....	27
5.	Statistik.....	28
<b>IV.</b>	<b>PUBLIZIERTE STUDIENERGEBNISSE .....</b>	<b>29</b>
1.	Veröffentlichung Piller et al. 2020 .....	29
<b>V.</b>	<b>ERWEITERTE ERGEBNISSE.....</b>	<b>43</b>
1.	Ergebnisse der Stichprobenberechnung .....	43
2.	Vergleich Kameranote mit der manuell gemessenen relativen Flächenveränderung.....	45
3.	Vergleich des verbesserten Kameranotes mit der manuell gemessenen relativen Flächenveränderung.....	48
4.	Relation zwischen Kameranote und Makronote.....	49
5.	Cut-Off-Werte des Kameranotes für den Makronote .....	52
6.	Relation zwischen Kameranote und Mikronote.....	55
<b>VI.</b>	<b>ERWEITERTE DISKUSSION.....</b>	<b>57</b>
1.	Wiedergabe der manuell gemessenen relativen Flächenveränderung durch den Kameranote .....	57
2.	Wiedergabe der manuell gemessenen relativen Flächenveränderung durch den überarbeiteten Kameranote.....	58
3.	Vergleich Makronote und Kameranote .....	59
4.	Vergleich des Kameranotes mit dem Makronote in Bezug auf die histologischen Ergebnisse .....	59
5.	Vor- und Nachteile einer kamerabasierten Erfassung am Schlachthof.....	60
6.	Vergleich der Ergebnisse für Footpad Dermatitis mit Ergebnissen bei	

	Hock Burn.....	61
7.	Ausblick .....	62
<b>VII.</b>	<b>ZUSAMMENFASSUNG .....</b>	<b>63</b>
<b>VIII.</b>	<b>SUMMARY .....</b>	<b>65</b>
<b>IX.</b>	<b>LITERATURVERZEICHNIS .....</b>	<b>67</b>
<b>X.</b>	<b>VERZEICHNIS DER ABBILDUNGEN UND TABELLEN .....</b>	<b>73</b>
1.	Abbildungsverzeichnis .....	73
2.	Tabellenverzeichnis .....	75
<b>XI.</b>	<b>ANHANG .....</b>	<b>77</b>
<b>XII.</b>	<b>DANKSAGUNG.....</b>	<b>79</b>

**ABKÜRZUNGSVERZEICHNIS**

<b>Abkürzung</b>	<b>Bezeichnung</b>
Accuracy	Genauigkeit (Erklärung siehe Anhang Tabelle I)
BLE	Bundesanstalt für Landwirtschaft und Ernährung
BMEL	Bundesministerium für Ernährung und Landwirtschaft
FPD	Foot-pad Dermatitis
KS	Kamerascore
Makroscore	Makroskopischer Score zur visuellen Beurteilung von FPD
Mikroscore	Mikroskopischer Score zur histopathologischen Beurteilung von FPD
NaN	Not a number / kein Wert kann zugeordnet werden
NPV	Negativer prädiktiver Wert (Erklärung siehe Anhang Tabelle I)
PPV	Positiver prädiktiver Wert (Erklärung siehe Anhang Tabelle I)
Sens	Sensitivität (Erklärung siehe Anhang Tabelle I)
Spec	Spezifität (Erklärung siehe Anhang Tabelle I)
TierSchNutzTV	Tierschutz-Nutztierhaltungsverordnung
Vgl.	vergleiche



## I. EINLEITUNG

Foot-pad Dermatitis (FPD, im Deutschen auch: Pododermatitis) ist eine entzündliche Veränderung der Fußballen beim Masthuhn (Martland, 1985). Aufgrund des häufigen Auftretens von FPD in der konventionellen Hühnermast (Saraiva et al., 2016) und dem Fakt, dass FPD als potentiell schmerzhaftes Veränderung angesehen wird (Algers und Berg, 2001) ist sie als Indikator für Tierwohl und Tiergesundheit gut geeignet (Allain et al., 2009) und kann reduziertes Wohlbefinden bei Masthühnern aufzeigen (Vanderhasselt et al., 2013). Mit der steigenden Bedeutung der Beurteilung von Tierschutzindikatoren im Rahmen der Umsetzung des §20 der Tierschutz-Nutztierhaltungsverordnung (2006) durch die Ausführungshinweise der Tierschutz-Nutztierhaltungsverordnung für Masthühner (2011) rückt in Deutschland auch immer mehr die automatisierte Erfassung von Tierwohlindikatoren in den Vordergrund. Durch standardisierte Erfassungsmethoden wird eine Vielzahl an Tieren bis hin zu ganzen Herden in kurzer Zeit beurteilt. Zudem ermöglichen standardisierte Systeme, welche Daten am Schlachthof erheben, eine sehr gute Vergleichbarkeit der erhobenen Befunde (De Jong et al., 2015).

Ziel dieser Studie war es, ein visuelles Bewertungsschema für Foot-pad Dermatitis sowie die Erfassung durch ein automatisiertes Kamerasystem am Schlachthof zu validieren und zu überprüfen, ob eine reine visuelle Beurteilung von FPD die histologisch messbare Tiefe der Entzündung sowie den histopathologischen Grad der Veränderung widerspiegeln kann. Hierzu sollten auf Grundlage der histopathologischen Ergebnisse Grenzwerte (sogenannte „Cut-Off“-Werte) für das makroskopische Scoringssystem berechnet und validiert werden.



## II. ERWEITERTE LITERATURÜBERSICHT

### 1. Pathologie und Pathogenese der Foot-pad Dermatitis

Foot-pad Dermatitis ist eine entzündliche, meist nekrotische Veränderung an den Fußballenflächen und Zehen (Shepherd und Fairchild, 2010) bei Masthühnern oder andern Geflügelarten wie z.B. Puten (Martland, 1984) oder Enten (Qin et al., 2019). FPD wurde zum ersten Mal in den 1980er Jahren beschrieben (Greene et al., 1985; Martland, 1985). Neben Fußballenentzündungen treten auch andere, ähnliche Hautveränderungen wie entzündliche Veränderungen der Fersenhöcker („Hock Burn“) und entzündliche Veränderungen im Brustbereich („Breast Burn“) auf (Martland, 1985; Haslam et al., 2007). Neben den negativen Auswirkungen auf das Tierwohl, wird die Tiergesundheit im Allgemeinen, die Laufaktivität, die Futteraufnahme sowie die Wachstumsrate durch FPD negativ beeinträchtigt (Qin et al., 2019).

FPD ist eine progressive Erkrankung der Fußballenhaut (Martland, 1985) und eine multifaktorielle Erkrankung (Oliveira et al., 2017), welche häufig bei Masthühnern auftritt (Saraiva et al., 2016). Genetik (Ask, 2010), Fütterung (Martland, 1985), jahreszeitlicher Einfluss (Kyvsgaard et al., 2013), Besatzdichte (Bruce et al., 1990) und vor allem feuchte Einstreu (Martland, 1985; Meluzzi et al., 2008) in Kombination mit chemischen Reizstoffen wie Ammoniak (Kyvsgaard et al., 2013) sind als Ursachen bereits identifiziert worden. Die FPD zeigt sich mit milden Verläufen durch Hautverfärbungen und Hyperkeratosen, kann jedoch auch schwere Verlaufsformen mit Schwellungen sowie Erosionen oder Ulzerationen annehmen (Martland, 1985).



## 2. Makroskopische Scoringssysteme für Foot-pad Dermatitis

FPD wurde als erstes in den 80er Jahren von Greene et al. (1985) und Martland (1985) beschrieben. Seitdem wurden viele Studien durchgeführt und visuelle Scoringssysteme zu FPD entwickelt, in den meisten Fällen werden dabei drei- oder vier-stufige Scores verwendet.

Martland (1985) beschreibt einen vierstufigen Score zur Beurteilung von FPD, dieser ist in der Tabelle 1 dargestellt. Martrenchar et al. (2002) beschreibt dagegen einen vierstufigen Score (Tabelle 2). In beiden Scoring Systemen wird die Größe der Läsion nach prozentualem Anteil veränderter Fläche des Fußballens bemessen.

Viele Studien beschreiben auch einen dreistufigen Score (Ekstrand et al., 1997; Bilgili et al., 2006; Welfare Quality® Assessment protocol for poultry (2009). Ekstrand et al. (1997) veröffentlichte einen dreistufigen Score und fasste damit den vorher im Jahre 1994 von Ekstrand und Svedberg (Ekstrand, 1997) publizierten sechsstufigen Score wie folgt zusammen (Tabelle 3). Auffällig sind hierbei die Fachbegriffe aus der Histologie wie Erosion und Ulzeration, obwohl der Score für eine makroskopische Beurteilung herangezogen wurde. Ekstrand et al. (1997) geben keine Begründung für die Zusammenlegung der Scores an, jedoch stellt diese Zusammenlegung in drei Scores eine eindeutige Vereinfachung mit mehr Praxisrelevanz dar.

Erosion sind ein umschriebener, oberflächlicher Gewebeverlust der Haut bis an das *Stratum Germinativum* (Pschyrembel, 2002), hingegen sind Ulzerationen Substanzdefekte der Haut/ Verlust von Epidermis mit darunterliegendem Entzündungsgewebe sowie Beschädigung des *Stratum Basale* (Michel et al., 2012).

Auch Bilgili et al. (2006) bewerteten FPD nach einem dreistufigen Score, wobei die Läsionen in diesem Scoring System nach der Größe in mm eingeteilt werden (Tabelle 4).

**Tabelle 1:** Makroskopischer Score für Foot-Pad Dermatitis nach Martland (1985).

Stufe	Definition
0	keine Läsionen
1	kleines Schorfareal < 5% der Fußballenfläche
2	großes Schorfareal < 25% der Fußballenfläche
3	schwere, großflächig Schorf bedeckte Ulzeration.

**Tabelle 2:** Makroskopischer Score für Foot-Pad Dermatitis nach Martrenchar et al. (2002).

Stufe	Definition
0	keine Läsionen
1	Läsion auf <25% der Fußballenfläche
2	Läsion auf 25% bis zu 50% der Fußballenfläche
3	Läsionen >50% der Fußballenfläche.

**Tabelle 3:** Makroskopischer Score für Foot-Pad Dermatitis nach Ekstrand et al. (1997) verglichen mit makroskopischem Score für Foot-Pad Dermatitis nach Ekstrand und Svedberg (Ekstrand et al., 1997).

Kategorie nach Ekstrand et al. (1997)	Stufe nach Ekstrand und Svedberg (Ekstrand 1997)	Definition
A keine Läsion	1	keine sichtbare Läsion: glatte Epidermis, keine Verfärbungen
B milde Läsion	2	Papille mit Hyperkeratose aber keine Verfärbung
B milde Läsion	3	milde/oberflächliche Läsionen: Erosionen oder Verfärbungen in der Epidermis
B milde Läsion	4	milde/oberflächliche Läsionen und Papille: Hyperkeratose und Erosionen oder Verfärbungen in der Epidermis
C schwere Läsion	5	schwere Ulzerationen: Verfärbungen, Ulzeration und Anzeichen einer Entzündungsreaktion
C schwere Läsion	6	schwere Ulzerationen und Papille: Verfärbungen, Ulzeration und Anzeichen einer Entzündungsreaktion

**Tabelle 4:** Makroskopischer Score für Foot-Pad Dermatitis nach Bilgili et al. (2006).

Stufe	Definition
0	keine Läsion vorhanden,
1	milde Läsion ( $\leq 7,5$ mm)
2	hochgradige Läsion ( $>7,5$ mm).

Ebenso wird im Welfare Quality® Assessment protocol for poultry (2009) eine dreistufige Klassifizierung, der mit Beispielbildern versehenen, fünf Scorestufen für FPD vorgeschlagen, eine schriftliche Beschreibung der Scores fehlt jedoch. Die Klassifizierung erfolgt nach dem in Tabelle 5 dargestellten Schema.

Wie dargestellt, gibt es eine Vielzahl von makroskopischen Bewertungssystemen die von Autoren angewendet werden. Dies hat zur Folge, dass aktuell länderübergreifend kein einheitliches Scoringssystem verwendet wird (Heitmann et al., 2018).

**Tabelle 5:** Makroskopischer Score für Foot-Pad Dermatitis nach Welfare Quality® Assessment protocol for poultry (2009).

Stufe	Definition
A	kein Hinweis für FPD (Score 0)
B	minimale Hinweise für FPD (Score 1 und 2)
C	Hinweise für FPD (Score 3 und 4)



### 3. Histologische Scoringssysteme für Foot-pad Dermatitis

Deutlich weniger Autoren beschäftigten sich mit der histologischen Einteilung der FPD. Martland (1985), Greene et al. (1985), McKeegan (2010) und Michel et al. (2012) berücksichtigten FPD aus histopathologischer Sicht.

Martland publizierte 1985 als erster eine Mikrofotografie von einem histologischen Schnitt durch die Epidermis eines Masthuhnfußballens mit FPD in der Hämatoxylin-Eosin-Färbung. Ebenso veröffentlichten Greene et al. (1985) histologische Schnitte von FPD beim Masthuhn, jedoch gehen beide Autoren nicht näher auf eine Differenzierung nach Schweregrad ein. McKeegan (2010) beschreibt einen 5-stufigen Score für FPD (Tabelle 6). Hier ist anzumerken, dass McKeegan den gleichen Score auch zur Anwendung bei Hock Burn empfiehlt. In diesem Score wird stärker differenziert, so werden zum Beispiel Ulzerationen entweder als eine fortgeschrittene Läsion mit Ulzeration in Stufe D und als eine fortgeschrittene Läsion mit extremer Ulzeration in Stufe E unterteilt. Erosion sind ein umschriebener, oberflächlicher Gewebeverlust der Haut bis an das *Stratum Germinativum* (Pschyrembel, 2002), hingegen sind Ulzerationen Substanzdefekte der Haut und Verlust von Epidermis mit darunterliegendem Entzündungsgewebe sowie Beschädigung des *Stratum Basale* (Michel et al., 2012).

**Tabelle 6:** Mikroskopischer Score für Foot-Pad Dermatitis nach McKeegan (2010).

Stufe	Definition
A	unbeeinträchtigte Haut
B	beginnende Läsion mit milder Hyperkeratose und milder Unterbrechung der Keratinschicht in der Hautschuppe
C	mittlere Läsion mit Acanthosis, erhöhter Vaskularisation und Hyperkeratose
D	fortgeschrittene Läsion mit Ulzeration, serozellulären Verkrustungen und Exsudation von Heterophilen
E	fortgeschrittene Läsion mit extremen Ulzerationen serozellulären Verkrustungen und Granulationsgewebe

Neben Hühnerfüßen mit physiologischer Haut und Fußballen mit Narben oder verheilten Läsionen differenziert Michel et al. (2012) drei verschiedene Arten von Läsionen und beschreibt diese histopathologisch (Tabelle 7). Im Vergleich zu McKeegan (2010) werden hier zum Beispiel Ulzerationen nur als eine Stufe betrachtet und nicht weiter unterteilt.

**Tabelle 7:** Mikroskopischer Score für Foot-Pad Dermatitis nach Michel et al. (2012).

Stufe	Definition
Typ I	milde epidermale Hyperplasie und/oder Hyperkeratose
Typ II	deutliche Hyperkeratose und Hyperplasie der Epidermis mit deutlicher dermalen Entzündungsreaktion
Typ III	Ulzerationen einhergehend mit Substanzverlust in der Epidermis

#### 4. Gegenwärtige Erfassung von Foot-pad Dermatitis in Deutschland

Momentan gibt es in Deutschland in den Ausführungshinweisen der Tierschutz-Nutztierhaltungsverordnung für Masthühner (2011) einen makroskopischen Score, der für die Bewertung für FPD herangezogen wird (Tabelle 8). Hier werden in Stufe 0 neben gesunden Füßen auch Füße mit leichter Hyperkeratose und Narben mitaufgenommen. Stufe 1 werden als oberflächliche Läsionen wie z.B. Erosionen klassifiziert, wobei im Gegensatz dazu schwere, tiefe Läsionen mit Ulzera als höchste Stufe (Stufe 2) eingeordnet werden.

Des Weiteren wurde durch den Runderlass des Niedersächsischen Ministeriums für Ernährung, Landwirtschaft und Verbraucherschutz (2014) ein weiteres Bewertungssystem für die Einstufung von FPD erlassen. Dieser wird in Niedersachsen angewendet und beinhaltet vier, anstatt der in den Ausführungshinweisen der Tierschutz-Nutztierhaltungsverordnung für Masthühner (2011) dargestellten drei Abstufungen (Tabelle 9).

**Tabelle 8:** Makroskopischer Score für Foot-Pad Dermatitis nach den Ausführungshinweisen der Tierschutz-Nutztierhaltungsverordnung für Masthühner (2011).

Stufe	Definition
0	Intakt (maximal leichte Hyperkeratose, keine Verfärbungen oder Narben)
1	Leichte, oberflächliche Läsionen (und Erosionen, Hyperkeratose und Verfärbungen der Fußballen)
2	Schwere, tiefe Läsionen (Ulzera und Narben, die sich entweder fast kreisförmig insbesondere im Bereich des Sohlenballens finden oder die Sohlen- und Zehenballen gleichermaßen betreffen)



**Tabelle 9:** Makroskopischer Score für Foot-Pad Dermatitis nach dem Runderlass des Niedersächsischen Ministeriums für Ernährung, Landwirtschaft und Verbraucherschutz (2014).

Stufe	Definition
0	Intakt
1	Leichte, oberflächliche Läsionen
2a	Mittlere, nur im Einzelfall tiefergehende Erosionen
2b	Schwere, tiefe Läsionen (kreisförmige Ulzera und Narben, die Sohlen- und Zehenballen gleichermaßen betreffen können)

Von 2013 bis 2015 wurden in ca. einem Drittel der EU bereits an Schlachthöfen die Prävalenz von FPD bei Masthühnern erfasst, was ca. 32% der geschlachteten Broiler der gesamten EU betrifft (EU-Generaldirektion Gesundheit und Lebensmittelsicherheit, 2016). Auch in Deutschland wird FPD von Amtstierärzten beurteilt, diese bewerten in der Regel 100 zufällig ausgewählte Füße einer Herde nach den Vorgaben der Tierschutz-Nutztierhaltungsverordnung visuell, wobei 50 Tiere zu Beginn und 50 Tiere am Ende der jeweiligen Herde begutachtet werden. In manchen Schlachthöfen werden bereits Kameras zur automatisierten Erfassung von FPD eingesetzt (Louton et al., 2018).

### **III. TIERE, MATERIAL UND METHODEN**

Die vorliegende Dissertation wurde im Rahmen des Forschungsprojektes „Verbundprojekt: Automatisierte Erfassung von Tierwohlindikatoren bei Geflügel (AutoWohl)“ erstellt. Das Projekt wurde vom Bundesministerium für Ernährung und Landwirtschaft (BMEL) über die Bundesanstalt für Landwirtschaft und Ernährung (BLE) gefördert (FKZ 2817903715). Wissenschaftliche Projektpartner waren der Fachbereich Nutztierethologie und Tierhaltung der Universität Kassel, die Fachgruppe Tierhaltung und Produkte (Studienschwerpunkt angewandte Geflügelwissenschaften) der Hochschule Osnabrück und das Institut für Tierhygiene, Tierschutz und Nutztierethologie, Stiftung Tierärztliche Hochschule Hannover.

Die wissenschaftliche Bearbeitung des Arbeitspaketes Masthuhn erfolgte durch den Lehrstuhl für Tierschutz, Verhaltenskunde, Tierhygiene und Tierhaltung der Ludwig-Maximilians-Universität München.





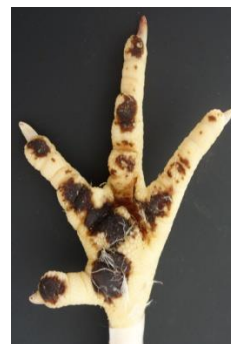
Die statistischen Berechnungen und Abbildungen für das Autowohlprojekt Masthuhn wurden in Zusammenarbeit mit Herrn Dr. Paul Schmidt mit der statistischen Programmiersprache R (R CORE TEAM, 2017) erstellt.

## 1. Beobachterabgleich

Vor Beginn der Studie wurde ein geeignetes Bewertungsschema zur Beurteilung von FPD ausgewählt. Das verwendete Bewertungsschema orientierte sich an einem modifizierten 5-stufigen Welfare Quality® Assessment protocol for poultry (2009) jedoch wurden die verschiedenen Score-Abstufungen genauer definiert (Tabelle 10). Für den Beobachterabgleich wurden 250 Masthühnerfüße zwischen Juli und November 2017 am Schlachthof „Donautal Geflügelspezialitäten, Zweigniederlassung der Lohmann & Co. KG“ in Bogen, Deutschland gesammelt und bei -18°C gelagert.

Eine möglichst gleichmäßige Verteilung der verschiedenen Scores wurde angestrebt. Die Proben wurden am 20. November 2017 aufgetaut und von fünf verschiedenen Beobachtern unabhängig und in zufälliger Reihenfolge nach obenstehenden Scoringsystem bewertet. Dem Beobachterabgleich lag eine überwiegend rein visuelle Beurteilung zu Grunde, jedoch war es den Teilnehmern erlaubt, bei Bedarf die Läsion oder den Fußballen mit einem „Precise PS 7215“ Messschieber (Burg-Wächter KG, Wetter, Deutschland) zu messen.

**Tabelle 10:** 5-stufiger Makroscore zur Bewertung von Foot-Pad Dermatitis modifiziert nach dem Welfare Quality® Assessment protocol for poultry (2009) (Bilder: Piller).

Score 0	Score 1	Score 2	Score 3	Score 4
				
Keine Läsion	Oberflächliche, zusammenhängende (einzelne) Läsion oder mehrere oberflächliche oder tiefe Läsionen in Summe bis 0,5 cm	Tiefe Läsionen über 0,5 cm bis zu 1,0 cm oder oberflächliche Läsion über 0,5 cm	Tiefe Läsionen über 1,0 cm	Läsion am Fußballen und eine oder mehrere tiefe Läsionen an den Zehen

## **2. Datenerhebung am Schlachthof**

### **2.1. Verwendete Tiere**

Insgesamt wurden 200 Füße von 39-42 Tage alten Ross 308 Masthühner am Schlachthof Donautal Geflügelspezialitäten, Zweigniederlassung der Lohmann & Co. KG in Bogen, Deutschland zwischen Februar und April 2018 entnommen. Die Tiere wurden nicht eigens für diese Studie, sondern zur Fleischgewinnung gezüchtet, gehalten und geschlachtet. Nur die bereits abgesetzten Füße wurden als Probenmaterial gesammelt, daher wurde kein Tier aufgrund dieser Studie verworfen.

### **2.2. Makroskopische und kamerabasierte Beurteilung**

In dieser Studie wurde zur makroskopischen Beurteilung von FPD ein modifiziertes fünfstufiges Schema in Anlehnung an das Welfare Quality® assessment protocol for poultry (2009) verwendet, welches vorab im Beobachtervergleich verglichen wurde (Tabelle 10).

Die 200 Füße wurden so gewählt, dass 40 Stück je makroskopischem Score gesammelt wurden. Diese wurden mit Hilfe eines Messschiebers „Precise PS 7215“ (Burg-Wächter KG, Wetter, Deutschland) manuell vermessen. Es wurde sowohl der Durchmesser des gesamten Fußballens sowie der Durchmesser der Fußballenläsion (falls vorhanden) gemessen. Es wurde in beiden Fällen immer an der breitesten Stelle gemessen. Der Sohlenballen beim Vogel (*Pulvinus metatarsalis*) ist in der Literatur als Bereich zwischen dem 2. und 4. Zehengrundgelenk definiert (Nickel et al., 1992). Nach der visuellen Beurteilung entsprechend dem obenstehenden Scoringschema wurde jeder Fuß mit einer „Sony Cyber-shot DSC-RX100“ Kamera (Sony Europe Limited, Surrey, UK) fotografiert. Anschließend wurden die nummerierten Füße in das stehende Schlachtband eingehängt. Die Füße wurden danach bei laufendem Schlachtband durch eine fest montierte „1.3 MP Farbkamera“ von IDS Imaging, Obersulm, Deutschland gefahren, welche mit Hilfe der Software „Chicken Check“ der CLK GmbH, Altenberge, Deutschland, und „Halcon“ von MVTec Software GmbH, München, Deutschland, Fotos der Füße erstellte. Die Software berechnet anschließend die prozentualen Flächenveränderungen von vorhandenen Läsionen. Der daraus berechnete Kamerascore gibt die prozentuale Veränderung zwischen gesamter Fußballenfläche und der Fläche der Läsion an. Es handelt sich hier um einen metrischen Wert, der positive Zahlen sowie 0 annehmen kann. Kurz nach der Messung wurde direkt am Messpunkt der Kamera die relative Luftfeuchtigkeit mit einem Thermo-Hygrometer „Testo 410-2“ (Testo AG, Lenzkirch, Deutschland) sowie die Beleuchtungsstärke mit einem Luxmeter „Pocket-Lux 2“ (LMT Lichtmesstechnik GmbH, Berlin, Deutschland) gemessen.



**Abbildung 1:** Beispielfoto ohne Softwareraster von Füßen, welche in der Schlachtkette durch die festinstallierte Kamera fahren (Bild: CLK GmbH).



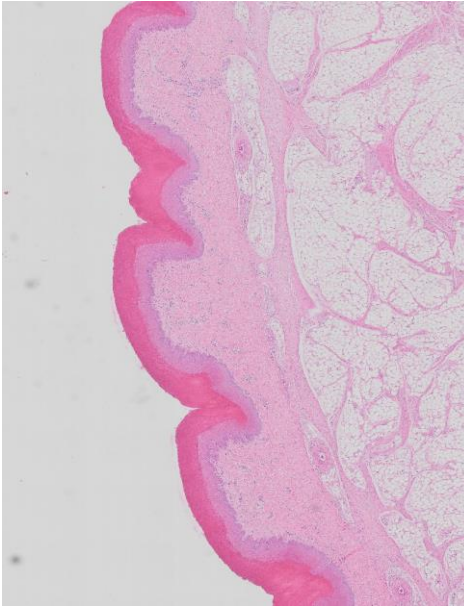
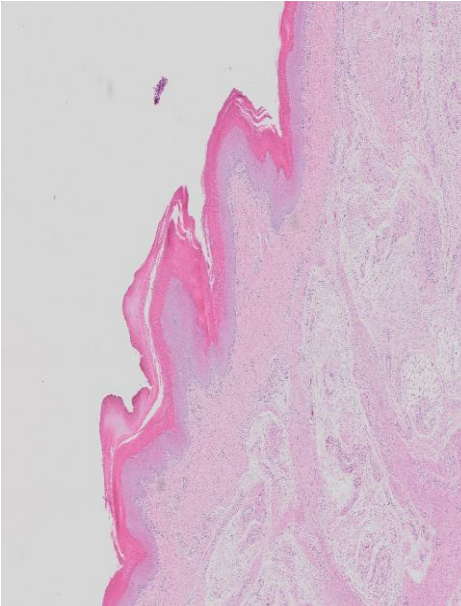
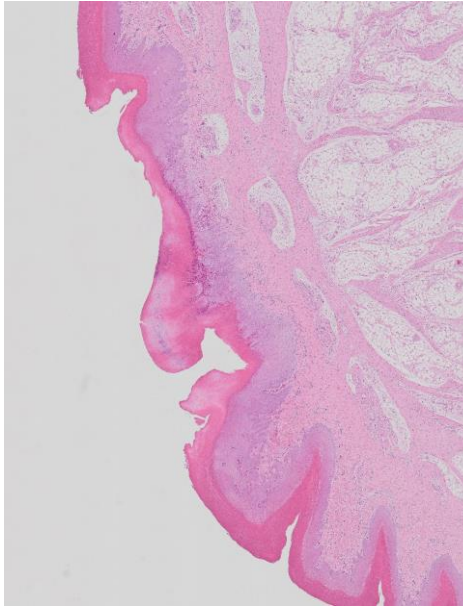
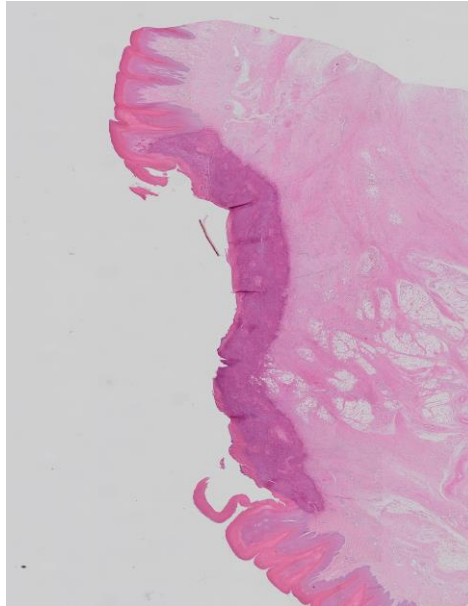
**Abbildung 2:** Beispielfoto mit Softwareraster von Füßen, welche in der Schlachtkette durch die festinstallierte Kamera fahren; grün = Fläche gesamter Fußballen, rot = Fläche der Läsion (Bild: CLK GmbH).

### 3. Histologische Beurteilung

Die Hühnerfüße wurden anschließend vom Schlachtband entnommen und zur histologischen Untersuchung an die Veterinärpathologie des Tiergesundheitsdienstes Bayern e.V. gebracht. Dort wurden die Proben in 10% gepufferter Formalinlösung für 24 Stunden fixiert und dann am Punkt der optisch größten Läsionsstelle mit einer Dicke von 4 µm geschnitten, mit Hämatoxylin-Eosin gefärbt, auf Glass-Objektträger fixiert und anschließend nach einem modifizierten Schema in Anlehnung an die Einteilung von Michel et al. (2012) von drei Veterinärpathologen beurteilt. Jede Messung wurde eindeutig einem Fuß zugeordnet, jedoch erfolgte die makroskopische und mikroskopische Beurteilung separat, sodass eine Beeinflussung durch vorherige gegebene Scores ausgeschlossen wurde.

Im Vorfeld fand ein Beobachterabgleich unter den Veterinärpathologen statt. Anschließend wurden die histologischen Schnitte mit einem „Olympus BX51“ Mikroskop (Olympus, Tokio, Japan) digitalisiert und die Dicke der physiologischen Haut vom *Stratum Corneum* bis zum *Stratum Basale*, sowie die Tiefe der Entzündung (falls vorhanden) vom *Stratum Corneum* bis zur Demarkationslinie der Entzündungsreaktion mit dem Programm „Olympus VS-ASW“ (Olympus, Tokio, Japan) vermessen.

**Tabelle 11:** Vier-stufiger Microscore zur Bewertung von Foot-Pad Dermatitis modifiziert nach dem Schema von Michel et al. (2012) (Bilder: Piller et al., 2020).

Score 0	Score 1	Score 2	Score 3
			
Physiologische Struktur der Epidermis	Leichte bis mittelgradige Hyperplasie und/oder Hyperkeratose der Epidermis; oberflächliche Stauung und Ödembildung; leichte Infiltration von Entzündungszellen; Basalmembran intakt	Deutliche Hyperplasie und Hyperkeratose der Epidermis; reichliche heterophile epidermale Exozytose; Infiltration von Entzündungszellen in der oberflächlichen Dermis; Basalmembran intakt	Ulzeration (Nekrose über alle Schichten der Epidermis; ggf. darunter Granulationsgewebe); Basalmembran nicht mehr intakt

#### **4. Stichprobenberechnung**

Vor Beginn der Studie wurden eine Stichprobenberechnung durchgeführt, da noch keine Werte über die mögliche Verteilung von den verschiedenen Makro- und Mikroscores vorhanden waren, wurden hypothetische Zuordnungsmatrizen für die Kategorien des Makro- und Mikroscores bestimmt, um Schätzungen für die erwarteten Wahrscheinlichkeiten generieren zu können.

Das eigentliche Ziel ist es, durch die so erhaltenen Daten die Wahrscheinlichkeiten und damit die Zuordnungsmatrix für die Kategorien des Mikro- und Makroscores zu schätzen. Durch das Fehlen der Datengrundlage wurden drei verschiedene Szenarios von Zuordnungsmatrizen betrachtet.



## 5. Statistik

Um die Übereinstimmung zwischen Mikroscore, Makroscore und Kameronascore zu berechnen wurden alle bedingten Wahrscheinlichkeiten der jeweiligen Score-Abstufungen für die jeweiligen Abstufungen des anderen Scores durch multinominale Regressionsmodelle geschätzt. Um die Relation zwischen Makroscore und der Tiefe der Entzündung zu bestimmen, wurde ein Hurtle-Gamma-Regressions-Modell erstellt. Für die Berechnung der Relation und des möglichen Bias (systemische Abweichung) zwischen manuell gemessener Flächenveränderung der FPD und des Kameronascore wurden Bland-Altman Plots nach Bland und Altman (1986) berechnet. In medizinischen Fachkreisen hat sich die Herangehensweise nach Bland und Altman namens „*Limits of agreement*“ für solche Fragestellungen als Standard etabliert. Wird auf dieser Art der Berechnung ein Bias entdeckt, kann mit Hilfe einer Korrektur eingegriffen werden. Diese Korrektur wird in zwei Schritten durchgeführt, da dies für den Vergleich zweier metrischer Merkmale notwendig ist. Im ersten Schritt wurde überprüft, ob die beiden Merkmale systematisch voneinander abweichen (Bias). Wird ein solcher Bias entdeckt, so könnten die Daten dahingehend korrigiert werden, da man eine solche Korrektur auch in der Wirklichkeit sehr leicht implementieren kann. Anschließend werden die Differenzen der beiden Methoden auf geeignete Art und Weise grafisch dargestellt (Bland-Altman Plot) und die Grenzen des 95% Prognoseintervalls eingezeichnet. Diese Grenzen des Prognoseintervalls werden mit den Grenzen verglichen, die aus theoretischer Sicht noch für eine Gleichheit der beiden Messungen sprechen würden („*Limits of agreement*“). Die „*Limits of agreement*“ sind unabhängig von den Ergebnissen theoretisch zu bewerten. Es stellt sich also die Frage, wie viel Abweichung zwischen den Methoden ist noch akzeptabel. Liegen die beobachteten Grenzen innerhalb dieser theoretischen Grenzen, so können wir auf eine Gleichheit der Methoden schließen, andernfalls nicht.

## **IV. PUBLIZIERTE STUDIENERGEBNISSE**

### **1. Veröffentlichung Piller et al. 2020**

A. Piller<sup>1</sup>, S. Bergmann<sup>1</sup>, A. Schwarzer<sup>1</sup>, M. Erhard<sup>1</sup>, J. Stracke<sup>2</sup>, B. Spindler<sup>2</sup>, N. Kemper<sup>2</sup>, P. Schmidt<sup>3</sup>, J. Bachmeier<sup>4</sup>, B. Schade<sup>5</sup>, B. Boehm<sup>5</sup>, E. Kappe<sup>5</sup>, H. Louton<sup>1\*</sup>

<sup>1</sup>Lehrstuhl für Tierschutz, Verhaltenskunde, Tierhygiene und Tierhaltung,  
Veterinärwissenschaftliches Department, Ludwig-Maximilians-Universität München,  
Veterinärstraße 13/R, 80539 München

<sup>2</sup>Institut für Tierhygiene, Tierschutz und Nutztierethologie, Stiftung Tierärztliche Hochschule  
Hannover, Bischofsholer Damm 15, 30173 Hannover

<sup>3</sup>Paul Schmidt, Statistisches Consulting, [www.statistische-modellierung.de](http://www.statistische-modellierung.de), Große Seestr. 8,  
13086 Berlin

<sup>4</sup>Tierärztliche Praxis, Grüner Weg 19, 94315 Straubing

<sup>5</sup>Bayerischer Tiergesundheitsdienst, Tierpathologie, Senator-Gerauer-Str. 23, 85586 Poing

**Validation of histological and visual scoring systems for foot-pad Dermatitis in broiler chickens**

**Validierung eines histologischen und visuellen Scoringsystems für Foot-pad Dermatitis bei Masthühnern**

Animal Welfare 2020, 29: 185-196

ISSN 0962-7286 / doi : 10.7120/09627286.29.2.185

**Angenommen am 15.12.2019**

**Impact Factor: 1.522 (2018/2019)**

© 2020 Universities Federation for Animal Welfare  
The Old School, Brevhouse Hill, Wheathampstead,  
Hertfordshire AL4 8AN, UK  
www.ufaw.org.uk

Animal Welfare 2020, 29: 185-196  
ISSN 0962-7286  
doi: 10.7120/09627286.29.2.185

## Validation of histological and visual scoring systems for foot-pad dermatitis in broiler chickens

A Filler<sup>†</sup>, S Bergman<sup>†</sup>, A Schwarzer<sup>†</sup>, M Erhard<sup>†</sup>, J Stracke<sup>‡</sup>, B Spindler<sup>‡</sup>, N Kemper<sup>‡</sup>,  
P Schmidt<sup>§</sup>, J Bachmeier<sup>§</sup>, B Schade<sup>¶</sup>, B Boehm<sup>¶</sup>, E Kappel<sup>¶</sup> and H Louton<sup>\*\*</sup>

<sup>†</sup> Department of Veterinary Sciences, Chair of Animal Welfare, Animal Behaviour, Animal Hygiene and Animal Husbandry, Faculty of Veterinary Medicine, LMU Munich, Veterinärstraße 13/R, 80539 Munich, Germany

<sup>‡</sup> Institute for Animal Hygiene, Animal Welfare and Farm Animal Behaviour, University of Veterinary Medicine Hannover, Foundation, Bischofsholer Damm 15, 30173 Hannover, Germany

<sup>§</sup> Paul Schmidt, Statistical Consulting for Science and Research, Große Seestr. 8, 13086 Berlin, Germany

<sup>¶</sup> Veterinary Practice, Grüner Weg 19, 94315 Straubing, Germany

<sup>¶</sup> Bavarian Animal Health Service, Department of Pathology, Senator-Gerauer-Str 23, 85586 Poing, Germany

\* Contact for correspondence: h.louton@lmu.de

### Abstract

In this study, an appropriate visual scoring system for foot-pad dermatitis was validated, considering the histologically measured depth of the inflammation zone and the histopathological grade (no lesion, mild lesion, ulcer). The aim being to evaluate whether the visual, macroscopic scoring of foot-pad dermatitis can represent the histological, microscopic findings. Two hundred Ross 308 broiler chicken feet (birds aged 39–42 fattening days) were collected at a slaughterhouse and scored macroscopically according to a modified version of the Welfare Quality® Assessment Protocol for Poultry. Afterwards, 200 histological slides (one per foot) were prepared, the extent of the inflammation measured and all slides scored by veterinarian pathologists using Michel *et al*'s modified scheme. The statistical relationship between microscopic and macroscopic score and depth of inflammation were estimated via regression models. Increasing macroscopic score was found to be linked with an increase in microscopic score and the depth of inflammation. In particular, feet without lesions and feet with ulcers were identifiable using the macroscopic score. Macroscopic scoring of foot-pad dermatitis can mirror histological findings once certain limitations are taken into account (superficial lesions were not clearly identifiable). Foot-pad dermatitis is considered a useful indicator of animal welfare and our findings suggest that visual, macroscopic scoring could be a practicable assessment tool.

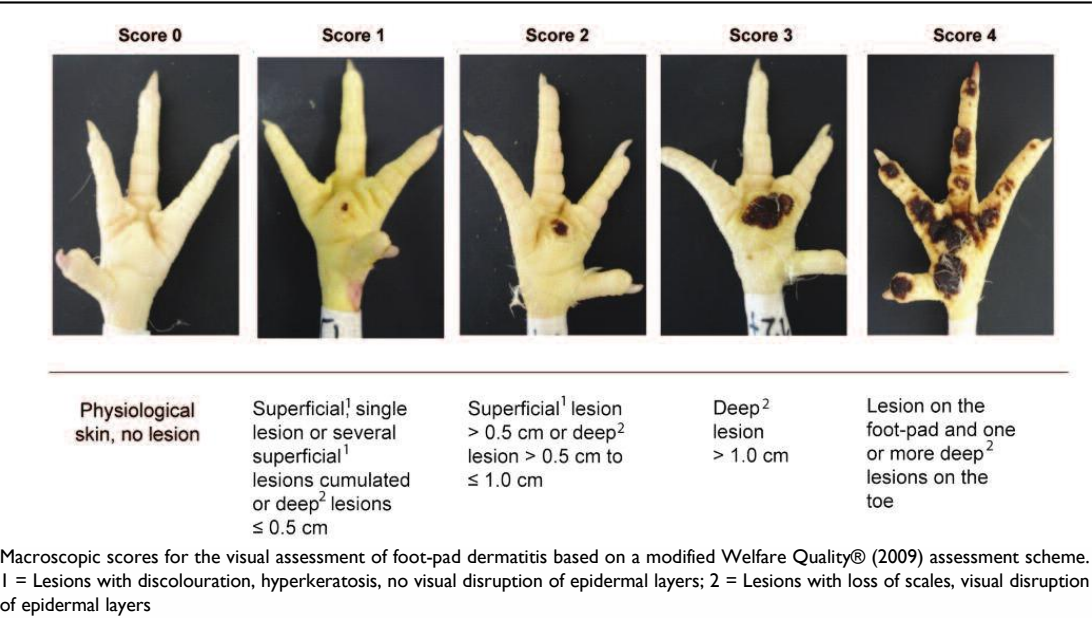
**Keywords:** animal welfare, animal welfare indicator, broiler, foot-pad dermatitis, histological validation, poultry

### Introduction

Foot-pad dermatitis (FPD) is a contact dermatitis on the plantar surface of the foot pad in broilers which can cause acute inflammatory and necrotic lesions (Greene *et al* 1985; Shepherd & Fairchild 2010). ‘Hock burn’ and ‘breast burn’ are similar forms of contact dermatitis, showing the same clinical signs in poultry as FPD but appearing on other parts of the body (Haslam *et al* 2007). The severity of FPD can depend on genetics (Ask 2010), stocking density (Bruce *et al* 1990; Spindler & Hartung 2011) and nutrition (Martland 1985), but the primary cause is wet litter (Martland 1985; Meluzzi *et al* 2008; Weber Wyneken *et al* 2015). Since FPD is a common occurrence in the conventional poultry industry (Saraiva *et al* 2016) and, almost certainly, painful (Algers & Berg 2001), it is a useful indicator of animal health and welfare (Allain *et al* 2009).

First described in the 1980s (Greene *et al* 1985; Martland 1985), this has been the subject of several studies throughout recent decades and several authors have had various scoring systems published that have sought to categorise FPD (Greene *et al* 1985; Martland 1985; Ekstrand *et al* 1997, 1998; Bilgili *et al* 2006; Welfare Quality® 2009; McKeegan 2010). As a result there is no fixed and uniform system currently in use (Heitmann *et al* 2018). The various systems differ in terms of the numbers of categories but, in most, a three- (Ekstrand *et al* 1998; Bilgili *et al* 2006; Welfare Quality® 2009) or a four-point scale (Martland 1985; Martrenchar *et al* 2002) is used to describe the macroscopic findings. Furthermore, Martland (1985), Greene *et al* (1985) and McKeegan (2010) considered the FPD lesions histopathologically, and Michel *et al* (2012) described FPD via a five-point scaled histological score, drawing a link to macroscopic results. The five-point

Figure 1



scaled histological scoring system for hock burn and FPD published by McKeegan (2010) contains the following lesion scores: (A) unaffected skin; (B) early lesion with mild hyperkeratosis and a mild disruption of keratin on the scale; (C) intermediate lesion with acanthosis, increased vascularity and hyperkeratosis; (D) advanced lesion with ulcer, serocellular crusting and exudation of heterophils; and (E) advanced lesion with extensive ulcer, serocellular crusting and granulation tissue.

Excluding normal feet and foot pads with scars or healing lesions, three types of lesions were described histologically by Michel *et al* (2012): type I, mild epidermal hyperplasia and/or hyperkeratosis; type II, marked hyperkeratosis and hyperplasia of the epidermis with marked dermal inflammation; and type III, ulceration with loss of epidermis. In Germany, there is currently an officially recommended system for visual scoring of FPD macroscopically: the German Order on the Protection of Animals and the Keeping of Production Animals (2006). Additionally, a modified version of this, implementing a four- instead of a three-point scale, was launched by the Implementary Rules of the Lower Saxonian Ministry of Nutrition (2015) and is in use in Lower Saxony, Germany. FPD evaluation in German slaughterhouses is randomly monitored by veterinarians who visually score 100 feet per flock. In some plants, cameras are used to scan for FPD (Louton *et al* 2018). From 2013 to 2015, 32% of EU broiler production, in approximately one-third of member states, were scoring FPD (European Commission Directorate-General for Health and Food Safety 2016).

The aim of this study was to determine whether a link exists between histological (ie microscopic) and visual (ie macroscopic) FPD scores. Furthermore, the relationship between these scores and the histologically measured depth of inflammation was assessed.

Materials and methods

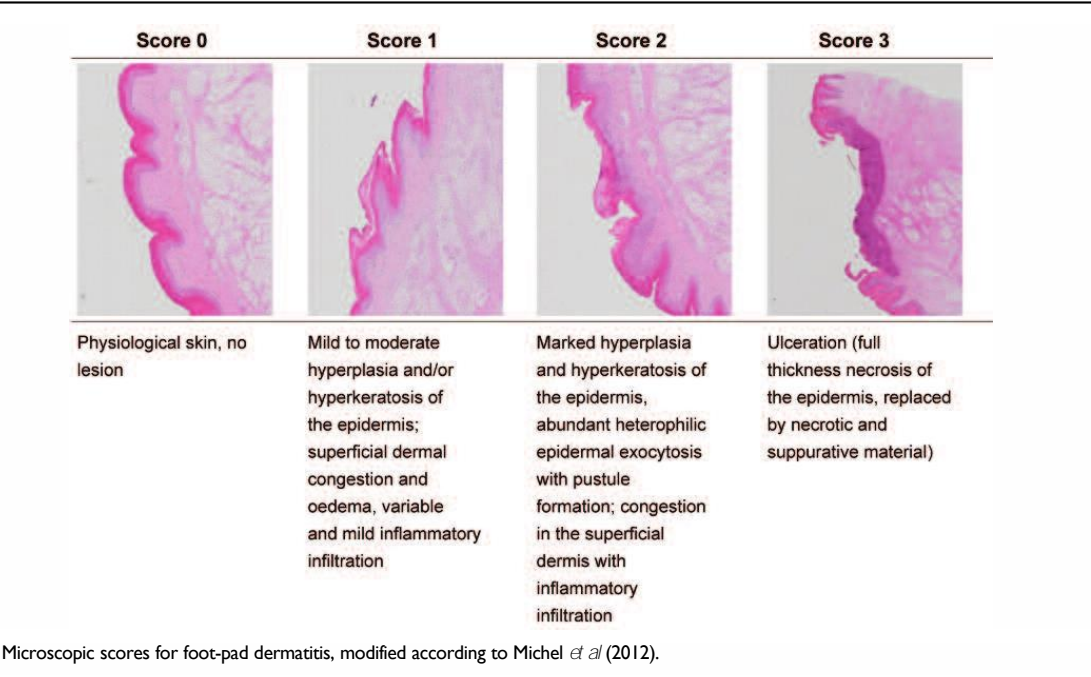
Initially, a scoring system to categorise FPD was chosen and inter-observer reliability was assessed. Then, broiler feet were collected and photographed, before a veterinarian measured the size of the lesions and scored feet. Histological slides were then produced allowing inflammation and foot lesions to be scored and measured.

All examined feet were from Ross 308 broilers aged 39–42 fattening days and taken from the same broiler slaughterhouse: Donautal Geflügelspezialitäten Zweigniederlassung der Lohmann & Co. KG in Bogen, Germany.

Inter-observer reliability

The aim was to choose an easy-to-use assessment scheme with a high inter-observer reliability for categorising of FPD scores. After an evaluation of pre-existing schemes, it was decided to use a modified version (Figure 1) of the Welfare Quality® Assessment Protocol for Poultry (2009). To validate the inter-observer reliability, 250 broiler feet were collected from the slaughterhouse between July and November 2017. Feet were stored at −18°C prior to use. The intention was to distribute scores as equally as possible and, after the samples were defrosted, five observers scored them independently and in a random order. Although assessing feet visually, observers were invited to use a

Figure 2



Precise PS 7215 digital measuring stick (Burg-Wächter KG, Wetter, Germany). Inter-observer reliability was calculated using the prevalence- and bias-adjusted kappa (PABAK/kappa<sub>pb</sub>) (Byrt *et al* 1993). According to Gunnarsson *et al* (2000), the following modified PABAK should be used when observers can assign more than two categories (k = number of categories; p<sub>0</sub> = relation between observed agreements):

$$\text{PABAK} = (\text{kp}_0 - 1) / (\text{k} - 1)$$

In addition, three veterinary pathologists from the Department of Pathology, Bavarian Animal Health Service, Poing, Germany, who scored all of the feet histologically, performed an inter-observer comparison with 20 randomly picked FPD slides and scored these individually. The PABAK values and the inter-observer reliabilities were calculated as per the first observer comparison.

**Data collection, applied macroscopic and applied microscopic scores**

Next, 200 feet (40 per macroscopic score) were obtained from the slaughterhouse and macroscopically scored according to Figure 1. Each foot was photographed (Sony Cyber-shot DSC-RX100 camera, Sony Europe Limited, Surrey, UK) and total foot pad and FPD sizes were measured directly on each foot, using the length between the most distant points. The feet were then taken to the Department of Pathology, Bavarian Animal Health Service,

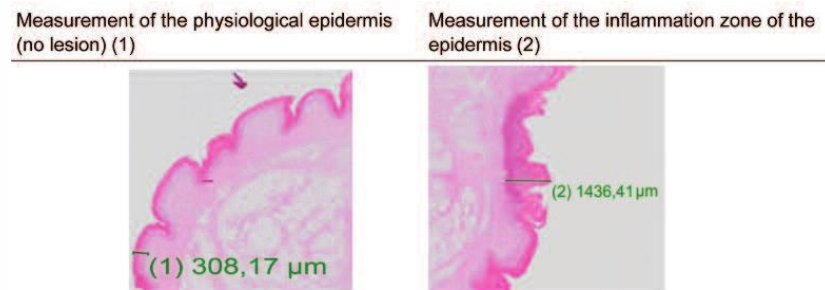
Poing, Germany and cut at the most critical point as determined visually. Samples were fixed in 10% buffered formalin for 24 h and then embedded in paraffin wax. Sections of 4-µm thickness were mounted onto glass slides, stained with haematoxylin and eosin, examined histologically and scored by veterinary pathologists using a modified version of Michel *et al*'s scheme (2012) (Figure 2). Additionally, each foot was examined histologically for potential pre-existing tendon inflammation.

Histological samples were digitalised by using an Olympus BX51 microscope (Olympus, Tokyo, Japan). Furthermore, physiological skin depth from stratum comeum to stratum basale and the depth of the inflammation zone from stratum comeum to the demarcation line (if present), were measured using the Olympus VS-ASW (Olympus, Tokyo, Japan) (Figure 3).

**Downscaled microscopic score**

In addition to the classification according to our modified four-point microscopic score (based on Michel *et al* 2012; Figure 2), a downscaled, three-point score was used simultaneously. For this, microscopic scores 1 and 2 were merged. Score 0 is identical in both schemes and represents feet with 'no lesion'. Score 1 in the downscaled microscopic score denotes 'mild lesion' and includes our modified microscopic scores 1 and 2. Downscaled microscopic score 2 represents ulcers and is identical to our modified microscopic score 3.

Figure 3



Examples of the histological measurements for the depth of the inflammation zone of the epidermis.

Table 1 PABAK results of the analysis of inter-observer reliability for the macroscopic scoring of foot-pad dermatitis (five observers; n = 250 feet).

Compared observer	1/2	2/3	1/3	2/4	3/4	1/4	2/5	3/5	4/5	1/5	Average
Match	225	226	223	228	228	221	227	227	227	226	
No match	25	24	27	22	22	29	23	23	23	24	
%	0.90	0.90	0.89	0.91	0.91	0.88	0.91	0.91	0.91	0.90	0.90
PABAK	0.88	0.88	0.87	0.89	0.89	0.86	0.89	0.89	0.89	0.88	0.88

### Macroscopic score regarding depth

After data had been collected, two veterinarians, who had also taken part in the first inter-observer comparison, visually scored the 200 pictures of the study feet, assessing only the depth of the inflammation. Scoring took place without any prior knowledge of the original macroscopic score as per Figure 1. The macroscopic scoring system addressing the depth of inflammation was as follows: 0 = no lesion; 1 = superficial lesions with discolouration, hyperkeratosis, no visual disruption of epidermal layers; and 2 = deep lesions with loss of scales, visual disruption of all epidermal layers.

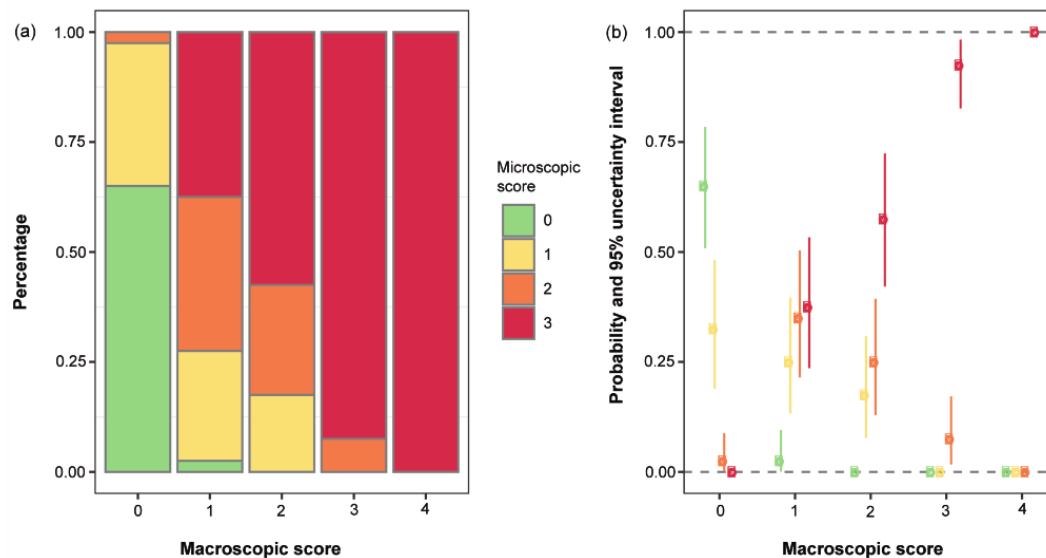
### Statistical analysis

To assess level of agreement between microscopic and macroscopic scores, the conditional probabilities of all microscopic categories given the macroscopic categories were estimated via multinomial regression models. To examine the relationship between macroscopic score and depth of inflammation, a Hurdle-Gamma regression model was used. Here, the Hurdle partly models the conditional probability for an inflammation depth  $> 0 \mu\text{m}$  given the macroscopic categories and the Gamma partly investigates the effect of the macroscopic score on the expected depth of inflammation  $> 0 \mu\text{m}$ . The Gamma distribution was chosen to account for the strictly positive and therefore skewed distribution of inflammation depth.

Prior to data collection, the required sample size was estimated. To this end, data for microscopic and macroscopic scores were simulated by assuming conditional probabilities of all microscopic categories given the macroscopic categories. Data were generated according to three different sets of conditional probabilities (optimistic, neutral and pessimistic) affecting the separability of score categories. Furthermore, equal sample size along the macroscopic categories was assumed. The generated datasets were analysed using the same multinomial regression model as above. Given the fitted model, it was checked if all hypothetically assumed conditional probabilities were within the estimated credible intervals. This step was repeated 10,000 times for each set of conditional probabilities and for sample sizes ranging from 20 to 100. The resulting proportion of simulations in which all conditional probabilities were estimated correctly was interpreted as statistical power.

All model parameters were estimated in a fully Bayesian way using the probabilistic programming language Stan (Carpenter *et al* 2017) and the wrapper package brms (Bürkner 2017) for the statistical programming language R (R Core Team 2018). For all estimated model parameters, point estimates as well as 95% credible/uncertainty intervals are reported. For the evaluation of cut-off values and the classification of attributes, the following performance measures were used: accuracy, sensitivity, specificity, positive and negative predictive value.

Figure 4



Showing (a) descriptive distribution of microscopic scores (0–3) for given macroscopic scores (0–4) of foot-pad dermatitis in broilers (Ross 308; 200 feet) and (b) probabilities and 95% uncertainty intervals of microscopic scores (0–3) for given macroscopic scores (0–4) of foot-pad dermatitis in broilers (Ross 308; 200 feet).

## Results

### Inter-observer reliability

As detailed in Table 1, the relations between matching of observed agreements varied from PABAK values of 0.86 up to 0.89, with an average of 0.88. The PABAK values of the second inter-observer comparison based on the histological findings were 0.80, 0.93 and 0.87, with an average of 0.87.

### Relation between macroscopic and microscopic score

Considering the macroscopic scores, lower or higher macroscopic scores corresponded, respectively, to lower or higher microscopic scores, although not on a one-to-one basis. Figure 4(a) shows the descriptive distribution of microscopic scores at given macroscopic FPD scores. The macroscopic scores contained mixed microscopic scores, with the exception of macroscopic score 4, which contained 100% microscopically ulcer (microscopic score 3).

Nearly all feet with microscopically no lesion (score 0) were visually diagnosed with no visual lesion (96.3%, macroscopic score 0), apart from one foot that was diagnosed with macroscopic score 1.

Figure 4(b) presents the results of the conditional probabilities of respective microscopic scores given macroscopic scores as well as the corresponding 95% credible/uncertainty intervals. For macroscopic score 0 (no visual lesion), microscopic score 0 (0.650, no lesion) had the highest probability. For macroscopic score 1, the probabilities for microscopic scores 1 (0.250), 2 (0.350) and 3 (0.375) were very

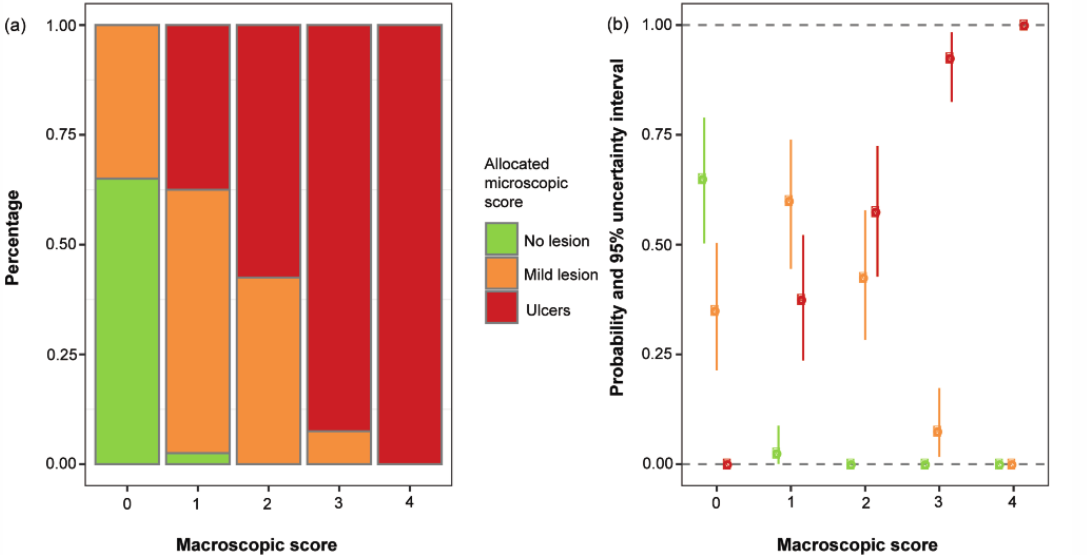
similar and microscopically no lesions were found with low probability (0.025). For macroscopic score 2, no probability for microscopic score 0, moderate probabilities for microscopic scores 1 (0.175) and 2 (0.250) and high probability for microscopic score 3 (0.575) were found. Macroscopic score 3 showed no probability for microscopic scores 0 and 1 (0.000), low probability for microscopic score 2 (0.075) and very high probability for microscopic score 3 (0.925, ulcer). Macroscopic score 4 only showed probability for microscopic score 3 (1.000, ulcer).

### Relation between macroscopic score and downscaled microscopic score

Figure 5(a) presents the relationship between macroscopic scores and downscaled microscopic scores (no lesion, mild lesion and ulcer). Macroscopic score 0 (macroscopically no lesion) contained 65.0% feet with microscopic no lesion and 35.0% feet with mild lesion. Macroscopic score 1 contained 2.5% feet with no lesions, 60% mild lesions and 37.5% ulcer. Macroscopic score 3 contained 7.5% mild lesion and 92.5% ulcer. Macroscopic score 4 contained only ulcer (100.0%).

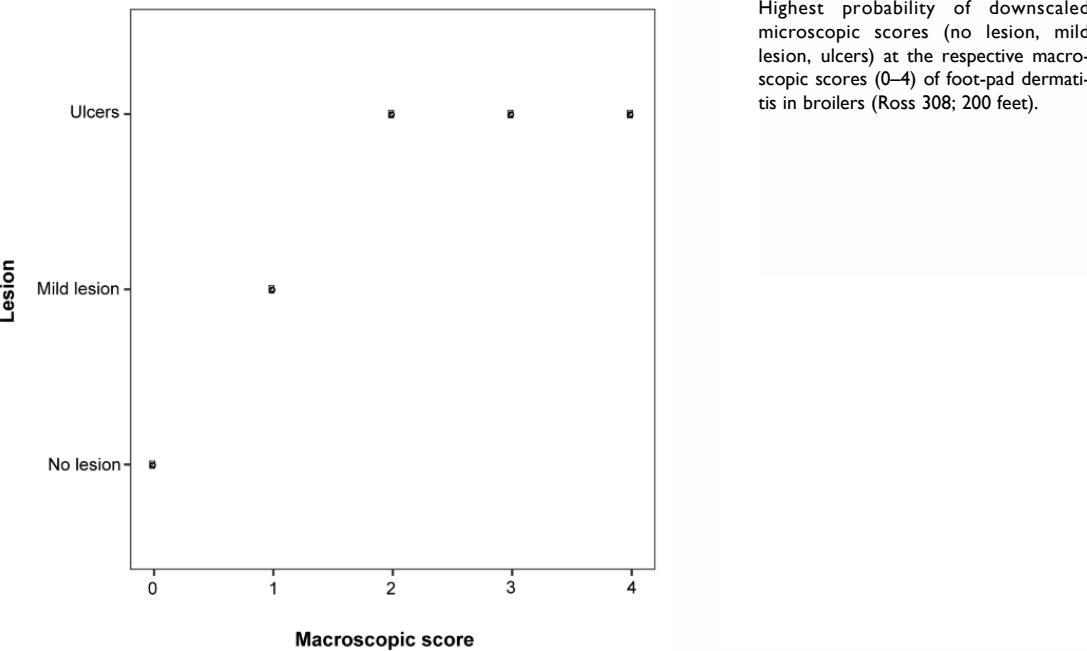
Nearly all feet with mild lesions were visually diagnosed with macroscopic score 1 (96.3%). The maximum peak for mild lesions was reached at macroscopic score 1. Along the increasing scale of macroscopic scores, a moderate decrease was seen in the prevalence of mild lesions (downscaled microscopic score 1) and a moderate increase in the prevalence of ulcer (downscaled microscopic score 2), although there was no one-to-one assignment to either.

Figure 5



Showing (a) descriptive distribution of downscaled microscopic scores (no lesion, mild lesion, ulcers) for given macroscopic scores (0–4) of foot-pad dermatitis in broilers (Ross 308; 200 feet) and (b) probabilities and 95% uncertainty intervals of downscaled microscopic scores (no lesion, mild lesion, ulcer) for given macroscopic scores (0–4) of foot-pad dermatitis in broilers (Ross 308; 200 feet).

Figure 6





## Histopathological validation of FPD in broiler chickens 191

Table 2 Performance measures of the macroscopic categories for the prediction of microscopic scores of foot-pad dermatitis in broilers (Ross 308; 200 feet).

(a)	Prediction				Performance			
	Macroscopic = 0	Macroscopic = 1	Macroscopic $\geq$ 2		Accuracy = 0.75			
Microscopic	No lesion	Mild lesion	Lesion with ulcer	Sens	Spec	PPV	NPV	
No lesion	26	1	0	0.96	0.92	0.65	0.99	
Mild lesion	14	24	20	0.41	0.89	0.60	0.79	
Lesion with ulcer	0	15	100	0.87	0.76	0.83	0.81	

(b)	Prediction				Performance			
	Macroscopic = 0	Macroscopic $\leq$ 2	Macroscopic > 2		Accuracy = 0.72			
Microscopic	No lesion	Mild lesion	Lesion with ulcer	Sens	Spec	PPV	NPV	
No lesion	26	1	0	0.96	0.92	0.65	0.99	
Mild lesion	14	41	3	0.71	0.73	0.51	0.86	
Lesion with ulcer	0	38	77	0.67	0.96	0.96	0.68	

Sens = Sensitivity: the probability that an animal belonging to category c is identified as category c;

Spec = Specificity: the probability that an animal not belonging to category c is identified as non-category c;

PPV = Positive predictive value: the probability that an animal is identified as category c and belongs to it;

NPV = Negative predictive value: the probability that an animal is not identified as category c and does not belong to category c;

Accuracy = Probability for the correct classification of the animals.

Figure 5(b) shows the conditional probabilities with corresponding uncertainty intervals of downscaled microscopic scores given the macroscopic scores. A probability for downscaled microscopic score 0 (no lesion) could only be found in feet with macroscopic scores 0 (0.650) and 1 (0.025). High probabilities of mild lesions were observed in feet with macroscopic score 1 (0.600) and 2 (0.425) and for ulcer in feet with macroscopic scores 2 (0.575), 3 (0.925) and 4 (1.000).

#### Cut-off data and resulting performance values

Figures 4(b) and 5(b) clearly indicate which microscopic or downscaled microscopic score had the highest probability for each macroscopic score in the given macroscopic score scheme. In Figure 6, the downscaled microscopic score with the highest probability is shown for the given macroscopic scores. At macroscopic score 0, no lesion (downscaled microscopic score 0) showed the highest probability. At macroscopic score 1, mild lesions (downscaled microscopic score 1) had the highest probability of diagnostic. For all other macroscopic scores (2, 3 and 4), ulcer (downscaled microscopic score 2) showed the highest probability. Thus, cut-off values can be calculated. The resulting performance values given a cut-off value for 'no lesion' at macroscopic score 0, 'mild lesion' at macroscopic score 1 and 'ulcer' at macroscopic score 2 can be seen in Table 2(a). In contrast, in Table 2(b), macroscopic scores 1 and 2 were categorised as microscopic 'mild lesion' and macroscopic score 3 and above as microscopic 'ulcer'. For the first calculated cut-off values (Table 2[a]), 26 out of 27 feet with no lesions were assigned to macroscopic score

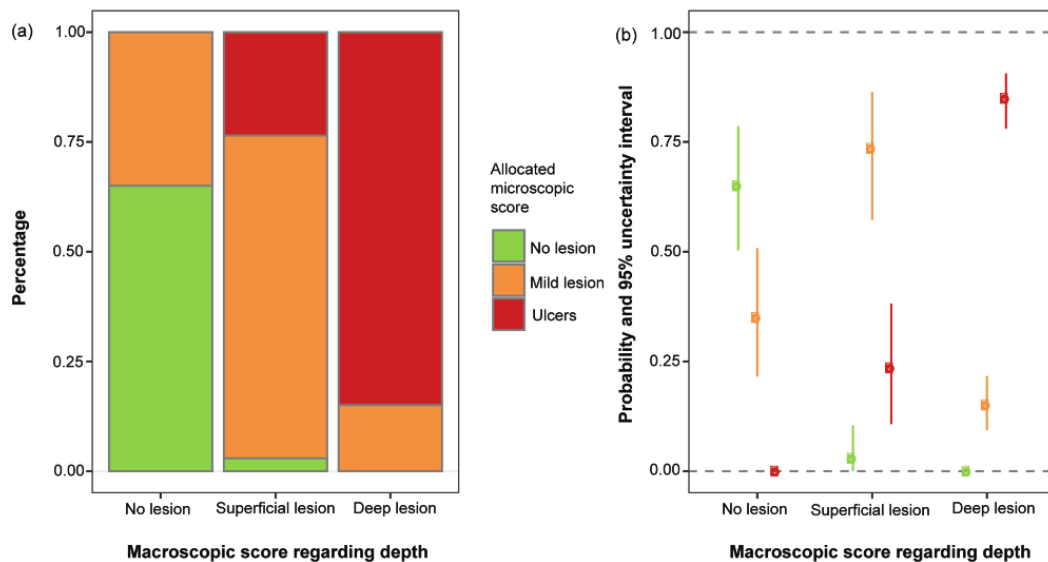
0 and thus could be diagnosed with a sensitivity of 0.96 and specificity of 0.92. Feet with mild lesions could be diagnosed with a sensitivity of 0.41 and specificity of 0.89. One hundred out of the 115 feet with ulcer were diagnosed with macroscopic score  $\geq$  2, resulting in a sensitivity of 0.87 and specificity 0.76. Given the cut-off value for ulcer at macroscopic score 2, the accuracy of all performance values was 0.75.

On the other hand, when macroscopic scores 1 and 2 were categorised as mild lesions resulting in a cut-off value for ulcer at macroscopic score 3 (Table 2[b]), all performance values were calculated with an accuracy of 0.72. Similarly to the results in Table 2(a), 26 out of 27 feet with no lesion were assigned to macroscopic score 0 and thus could be diagnosed with a sensitivity of 0.96 and specificity of 0.92. Feet with mild lesions could be predicted with a sensitivity of 0.71 and specificity of 0.73. Feet with ulcer could be predicted with a sensitivity of 0.67 and specificity of 0.96. In addition, the positive and negative predictive values were calculated for the diagnosis of the downscaled microscopic score (Table 2[a] and [b]).

#### Relation between macroscopic score regarding depth and downscaled microscopic score

Figure 7(a) shows the descriptive distribution of macroscopic scores considering the depth of inflammation in relation to the downscaled microscopic scores. Feet with no visible lesion (macroscopic score regarding depth 0) included mainly feet with no histological lesion (downscaled microscopic score 0), and ulcer were not detected. Mostly mild lesions were observed in feet with visually

Figure 7



Showing (a) descriptive relation between downscaled microscopic scores (no lesion, mild lesion, ulcers) and macroscopic scores considering the depth of inflammation (no lesion, superficial lesion, deep lesion) of foot-pad dermatitis in broilers (Ross 308; 200 feet) and (b) probabilities and 95% uncertainty intervals of downscaled microscopic scores (no lesion, mild lesion, ulcers) for given macroscopic scores regarding depth (no lesion, superficial lesion, deep lesion) of foot-pad dermatitis in broilers (Ross 308; 200 feet).

superficial lesions (macroscopic score regarding depth 1). Visually deep lesions (macroscopic score regarding depth 2) included mainly ulcer.

Figure 7(b) shows the different probabilities with 95% uncertainty intervals for the macroscopic scores regarding depth and the downscaled microscopic scores (no lesion, mild lesion, ulcer). Feet with no visual lesion had the highest probability for histological diagnosis with no lesion (0.650). Visual superficial lesions (macroscopic score regarding depth 1) had the highest probability for diagnosis with mild lesions (0.735) and the lowest for no lesions (0.003). Visually deep lesions (macroscopic score regarding depth 2) had very high probability for diagnosis with ulcer (0.850), low probability for mild lesions (0.150) and none for no lesions (0.000).

#### Relation between macroscopic score and histological depth of inflammation

The histologically measured depth of inflammation in lesions of the foot pad is presented for the different macroscopic scores (0, 1, 2, 3 and 4) in a boxplot (Figure 8[a]). An increase in the depth of inflammation with increasing macroscopic score up to macroscopic score 3 was found, with mean values increasing from 937.1 (macroscopic score 1) to 1,087.4 (macroscopic score 2) to 1,443.0  $\mu\text{m}$  (macroscopic score 3). The mean value for macroscopic score 0 was 8.0  $\mu\text{m}$ , and 97.5% of all feet assigned to macroscopic score 0 were measured

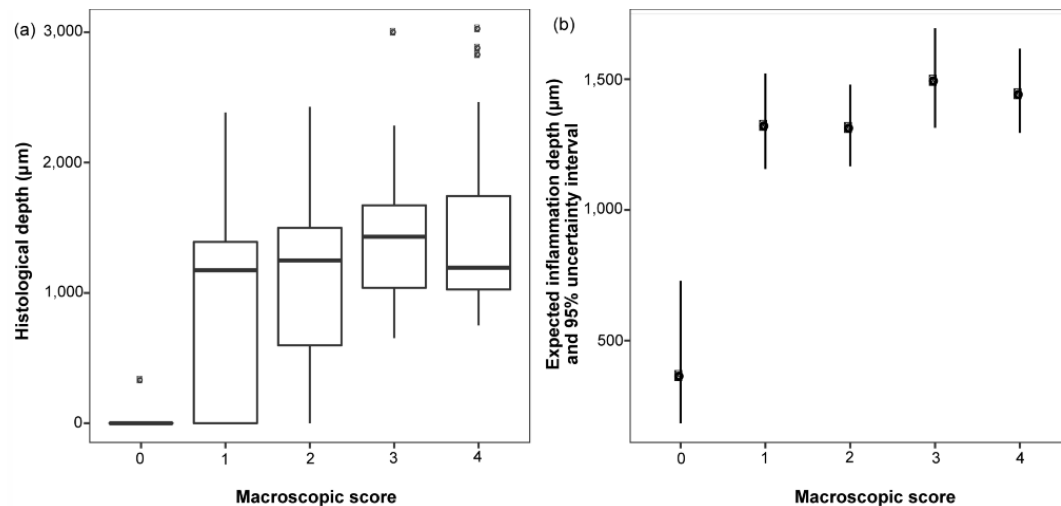
with 0.0  $\mu\text{m}$  (with no histological lesion). Feet assigned to macroscopic score 4 showed a mean inflammation depth of 1,422.2  $\mu\text{m}$ , slightly less than those assigned to macroscopic score 3. Macroscopic scores 1 and 2 included 27.5 and 12.5% feet, respectively, with 0.0- $\mu\text{m}$  depth of inflammation. A strong decrease in feet was noted with 0.0- $\mu\text{m}$  depth of inflammation with increasing macroscopic score up to macroscopic score 2. The boxplot in Figure 8(a) shows macroscopic scores 1 to 4 to have similar 0.50 quantiles and to vary in the space between 0.25 and 0.75 quantiles.

The relation between macroscopic scores and expected inflammation depth was further analysed by a regression model (Hurdle-Gamma). The estimated expected inflammation depth is shown in Figure 8(b). This analysis revealed significant differences between macroscopic score 0 and the remaining macroscopic scores (1 to 4). In all cases, macroscopic score 0 showed the significantly smallest expected inflammation depth. Differences between higher macroscopic scores could not be confirmed.

#### Tendinitis

In feet with microscopic scores 0, 1 and 2, no tendinitis was observed. The two feet (out of 200 examined feet) with tendinitis were scored as microscopic score 3 (ulcer). As both feet were categorised as macroscopic score 4, there was no inflammation of the tendons detected in macroscopic scores 0 to 3.

Figure 8



Showing (a) boxplot of relation between macroscopic scores (0–4) and the histologically measured depth (µm) of inflammation of foot-pad dermatitis in broilers (Ross 308; 200 feet) and (b) expected inflammation depth (µm) and 95% uncertainty intervals for respective macroscopic scores (0–4) of foot-pad dermatitis in broilers (Ross 308; 200 feet).

## Discussion

### Inter-observer reliability

The PABAK values of the inter-observer comparison for the rating of both the macroscopic and the histological findings ranged from 0.81 to 1.00. According to Landis and Koch (1977), this range represents an almost perfect reliability score.

### Macroscopic score mirroring histological findings with limitations

Ulcerations are defined as lesions with perforated and destroyed epidermis including stratum basale, whereas mild lesions show hyperkeratosis, hyperplasia or mild inflammation with intact stratum basale (Michel *et al* 2012). In this study, we used a visual, macroscopic scoring system for FPD with and without considering the depth of inflammation in the observed lesion and compared it with a histological, microscopic scoring. Our histological findings identified the macroscopic scores regarding depth as most suitable predictors of the histopathological presence of mild lesions or ulcers, taking various limitations into account. In general, the results illustrated a clear relationship between macroscopic scores and histopathological findings. For the macroscopic examinations, feet with no histologically detectable lesions were mainly assigned to feet with no visually detected lesions; similarly, feet with histologically identified ulcers were related to feet with visually scored bigger and deeper lesions.

In line with Michel *et al* (2012), our analyses revealed a reliable link between macroscopic and microscopic scoring with a few exceptions. Mayne *et al* (2006) showed similar results in turkeys and observed that feet with externally

normal foot pads revealed histopathological evidence of lesions. Our results confirm that histologically detectable mild lesions with hyperkeratosis and/or erosions can be distributed widely throughout feet with visually no detected lesions, mild/superficial lesions and big/deep lesions. Thus, macroscopic scores 1 and 2 (mild lesions) are not easy to distinguish visually by use of a macroscopic score from feet with no lesions or ulcers. This limitation may explain the poor performance values for the diagnosis of mild lesions when only a macroscopic score was received (ie when macroscopic score 2 was the cut-off value to assign ulcers). However, the diagnoses of feet with no lesions and feet with ulcers had high performance values in this scenario, and these feet could be assigned macroscopic scores with high probabilities.

In contrast, when macroscopic scores 1 and 2 were categorised as mild lesions (ie when macroscopic score 3 was the cut-off value to assign ulcers), we found a much higher sensitivity (0.71 instead of 0.41) and a lower specificity (0.73 instead of 0.89) for the diagnosis of mild lesions. Thus, mild lesions were more likely to be diagnosed correctly. However, lesions with ulcers (macroscopic scores 3 and 4) could then be diagnosed with a lower sensitivity (0.67 instead of 0.87) and a higher specificity (0.96 instead of 0.76) compared with the first presented cut-off value.

To summarise, if the focus is on detecting ulcers, high sensitivity (albeit low specificity) would be achieved by using the first cut-off value. The second cut-off value would allow a better detection of mild lesions. The accuracy for the first cut-off value (0.75) was higher than for the second (0.72).

The results of this study confirm the correlation between macroscopic and histological findings reported by

McKeegan (2010) and Michel *et al* (2012). In the study of McKeegan (2010), feet with visually scored lesions mainly had histologically detectable inflammatory lesions although a number of exceptions were noted. Similarly, we found that certain feet with no visual evidence of lesions showed mild lesions in the histological examination.

#### Importance of identifying ulcers

In the scoring system of the German Order on the Protection of Animals and the Keeping of Production Animals (2006), the highest score (2b) is defined as the presence of an ulcer. In addition, it is generally accepted that FPD, particularly in the form of severe ulcers, is likely to cause pain (Algers & Berg 2001). However, to date, no study has differentiated or categorised grades of FPD or inflammation depths in terms of different pain levels or been able to clearly define which grade or inflammation depth is relevant as regards animal welfare. Michel *et al* (2012) supported the hypothesis that the level of pain increases relative to the size of the lesion and according to the different levels of severity (depth of inflammation) seen between microscopic scores 2 and 3. The authors drew on the work of Arnould and Colin (2008) to formulate their conclusion and the enhanced reaction to manipulation with increasing severity of FPD. Hocking and Wu (2013) also adhered to the notion that FPD is likely to cause pain in turkeys. Moreover, McKeegan (2010) found chickens with FPD to move more slowly than those with no lesions and walk better with analgesia than without. Martland (1985) also drew a possible link between decreased appetite associated with pain and reduced growth rate. Furthermore, Weber Wyneken *et al* (2015) found a significant interaction between FPD and analgesia regarding stride length in turkeys and differences between birds with low- and high-grade FPD. However, lameness in broilers may not be related to FPD and can occur for other reasons, such as chondronecrosis and osteomyelitis (Wideman 2016).

Until scientific research confirms a difference in pain level according to different grades of FPD, an existing ulcer with inflammation should be considered painful (Algers & Berg 2001). Following this assumption, a macroscopic scoring system for FPD able to differentiate feet with no histological lesion from feet with ulcerations should be used to record FPD as an animal welfare indicator. The probability for FPD with an ulcer is highest at macroscopic scores 2 and above in our scoring system. Thus, a lesion with a diameter greater than 0.5 cm is likely to be a deep lesion with an ulcer and considered painful.

Overall, to diagnose an ulcer, we can recommend a simplified three-point macroscopic scoring system for FPD to differentiate between feet with no lesion, feet with mild lesions and feet with ulcers. Several, similar, three-point scores have been published, such as the summarised score of the Welfare Quality® Assessment Protocol for Poultry (2009) or that of Heitmann *et al* (2018). Considering, in particular, a camera-based and standardised detection of FPD, a three-point scale that can clearly distinguish an ulcer from other lesions would be applicable. Based on our results, a good suggestion for a useful score would be:

- Score 0 = no lesion;
- Score 1 = mild lesions (superficial lesions with discolouration, hyperkeratosis and no visual disruption of the epidermal layers, size  $\leq 0.5$  cm in diameter);
- Score 2 = deep lesions with loss of scales, visual disruption of all epidermal layers, ulcers and/or crusting (size  $> 0.5$  cm in diameter).

A similar scoring system has been published by Ekstrand *et al* (1998), however the authors did not define specific lesion sizes. In the commonly used scoring systems at slaughter, macroscopic scores are used to categorise FPD (Louton *et al* 2018), and according to our results, a routine histological examination is not necessary to score FPD effectively. Our study shows that it is possible to extrapolate from the macroscopic scoring of FPD to the type of lesion, with certain limitations.

#### Depth of inflammation and tendinitis

Although no significant statistical differences were found between the different macroscopic scores according to the histologically determined depth of inflammation, we observed an increasing inflammation depth with increasing macroscopic score. Mirroring the results of Heitmann *et al* (2018), a positive correlation was found between inflammation depth and size of the lesion in ulcers. However, since Heitmann *et al* (2018) measured the total thickness of skin layers (while we measured the depth of the inflammation zone), a one-to-one comparison of the results is not possible. Heitmann *et al* (2018) described large ulcers (90 mm<sup>2</sup>) with an increase in thickness of skin layers to 1,535.5  $\mu$ m, a value similar to the mean depth of the inflammation zone in macroscopic scores 3 (1,443.0  $\mu$ m) and 4 (1,422.2  $\mu$ m) measured in our study.

Another way of scoring FPD could be via the development of a practical and non-invasive technique for detecting the depth of inflammation at the slaughterhouse, eg by using infra-red cameras (Wilcox *et al* 2009). Such an approach might increase the likelihood that mild lesions would be diagnosed with greater accuracy.

Affected tendons were found only in individual cases and no significant relation was discovered between FPD and detectable inflammation of tendons.

#### Animal welfare implications

The validated scoring systems discussed here can contribute to the improvement of animal welfare within the poultry industry. Taking into account our results, we recommend a three-point, visual scoring system for the detection of ulcers in FPD. Moving forward, a macroscopic, comprehensive and standardised detection of FPD could be especially beneficial in improving the health and welfare of broiler chickens. Future research should evaluate if the results of this study can also be applied to an automatic scoring system with camera-based detection of FPD. Furthermore, more research is needed to consider the perception of pain at different inflammation depths or grades of FPD and additionally future research should focus on reducing the prevalence of FPD. Nevertheless, this histologically validated, veterinarian-endorsed confirmation of macroscopic scoring and inflammation depth of FPD lend validity to these macroscopic scores operating as indicators of animal welfare.

## Conclusion

In general, our study elaborates significant links between the examined macroscopic and microscopic scoring systems and demonstrates the extent of their comparability. However, limitations, notably the diagnosis of mild lesions such as hyperkeratosis, erosions and other superficial lesions need to be considered. A routine examination based upon visual scoring can be recommended for the diagnosis of ulcers; it accurately mirrors histopathological findings, especially when feet with no lesions and feet with ulcers are the main focus. Although no statistically significant differences in inflammation depth between the visually present lesions were found, an increase in inflammation depth with increasing severity of FPD was detected. A link between tendinitis and FPD was not found.

## Acknowledgements

This project was supported by funds from the Federal Ministry of Food and Agriculture (BMEL) based on a decision from the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme (grant number: FKZ 2817903715). The funding source played no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. We thank Dr Sven Reese for providing the Olympus BX51 microscope and great scientific support, Dr Verena Lietze for editing of the scientific language as well as Stefan and Yeliz Wiesbeck, Astrid Nagel and Birgit Isele-Rüegg for excellent technical assistance. There is no conflict of interest regarding any financial, personal or other relationships with other people or organisations that could inappropriately influence or be perceived to influence this study.

## References

- Algers B and Berg C 2001 Monitoring animal welfare on commercial broiler farms in Sweden. *Acta Agriculturae Scandinavica, Section A, Animal Science* 51: 88-92. <https://doi.org/10.1080/090647001316923135>
- Allain V, Mirabito L, Arnould C, Colas M, Le Bouquin S, Lupo C and Michel V 2009 Skin lesions in broiler chickens measured at the slaughterhouse: Relationships between lesions and between their prevalence and rearing factors. *British Poultry Science* 50: 407-417. <https://doi.org/10.1080/00071660903110901>
- Arnould C and Colin L 2008 Relation between a simple measure of activity, lameness and pododermatitis in meat type chickens. In: Koene P (ed) *Book of Abstracts of the 4th International Workshop on the Assessment of Animal Welfare at Farm and Group Level* p 73. 10-13 September 2008, Ghent, Belgium. Ponsen and Looijen: Wageningen, The Netherlands
- Ask B 2010 Genetic variation of contact dermatitis in broilers. *Poultry Science* 89: 866-875. <https://doi.org/10.3382/ps.2009-00496>
- Bilgili SF, Alley MA, Hess JB and Nagaraj M 2006 Influence of age and sex on footpad quality and yield in broiler chickens reared on low and high density diets. *The Journal of Applied Poultry Research* 15: 433-441. <https://doi.org/10.1093/japr/15.3.433>

- Bruce DW, McIlroy SG and Goodall EA 1990 Epidemiology of a contact dermatitis of broilers. *Avian Pathology* 19: 532-537. <https://doi.org/10.1080/03079459008418705>
- Bürkner PC 2017 brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80: 1-28. <https://doi.org/10.18637/jss.v080.i01>
- Byrt T, Bishop J and Carlin JB 1993 Bias, prevalence and kappa. *Journal of Clinical Epidemiology* 46: 423-429. [https://doi.org/10.1016/0895-4356\(93\)90018-V](https://doi.org/10.1016/0895-4356(93)90018-V)
- Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P and Riddell A 2017 Stan: A probabilistic programming language. *Journal of Statistical Software* 76: <https://doi.org/10.18637/jss.v076.i01>
- Ekstrand C, Algers B and Svedberg J 1997 Rearing conditions and foot-pad dermatitis in Swedish broiler chickens. *Preventive Veterinary Medicine* 31: 167-174. [https://doi.org/10.1016/S0167-5877\(96\)01145-2](https://doi.org/10.1016/S0167-5877(96)01145-2)
- Ekstrand C, Carpenter TE, Andersson I and Algers B 1998 Prevalence and control of foot-pad dermatitis in broilers in Sweden. *British Poultry Science* 39: 318-324. <https://doi.org/10.1080/00071669888845>
- European Commission Directorate-General for Health and Food Safety 2016 *Use of slaughterhouse data to monitor welfare of broilers on farm*. Luxembourg Publications Office of the European Union: Luxembourg
- German Order on the Protection of Animals and the Keeping of Production Animals 2006 *Amended and promulgated on August 22nd, 2006, last changed on February 5th, 2014*. Verordnung zum Schutz landwirtschaftlicher Nutztiere und anderer zur Erzeugung tierischer Produkte gehaltener Tiere bei ihrer Haltung: Tierschutz-Nutztierhaltungsverordnung (TierSchNutzV) in der Fassung der Bekanntmachung vom 22 August 2006 (BGBl. I S 2043), die durch Artikel 1 der Verordnung vom 5 Februar 2014 (BGBl. I S 94) geändert worden ist
- Greene JA, McCracken RM and Evans RT 1985 A contact dermatitis of broilers — clinical and pathological findings. *Avian Pathology* 14: 23-38. <https://doi.org/10.1080/03079458508436205>
- Gunnarsson S, Algers B and Svedberg J 2000 Description and evaluation of a scoring system of clinical health in laying hens. In: Gunnarsson S (ed) *Laying Hens in Loose Housing Systems*. Doctoral thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden
- Haslam SM, Knowles TG, Brown SN, Wilkins LJ, Kestin SC, Warriss PD and Nicol CJ 2007 Factors affecting the prevalence of foot pad dermatitis, hock burn and breast burn in broiler chicken. *British Poultry Science* 48: 264-275. <https://doi.org/10.1080/00071660701371341>
- Heitmann S, Stracke J, Petersen H, Spindler B and Kemper N 2018 First approach validating a scoring system for foot-pad dermatitis in broiler chickens developed for application in practice. *Preventive Veterinary Medicine* 154: 63-70. <https://doi.org/10.1016/j.prevetmed.2018.03.013>
- Hocking PM and Wu K 2013 Traditional and commercial turkeys show similar susceptibility to foot pad dermatitis and behavioural evidence of pain. *British Poultry Science* 54: 281-288. <https://doi.org/10.1080/00071668.2013.781265>

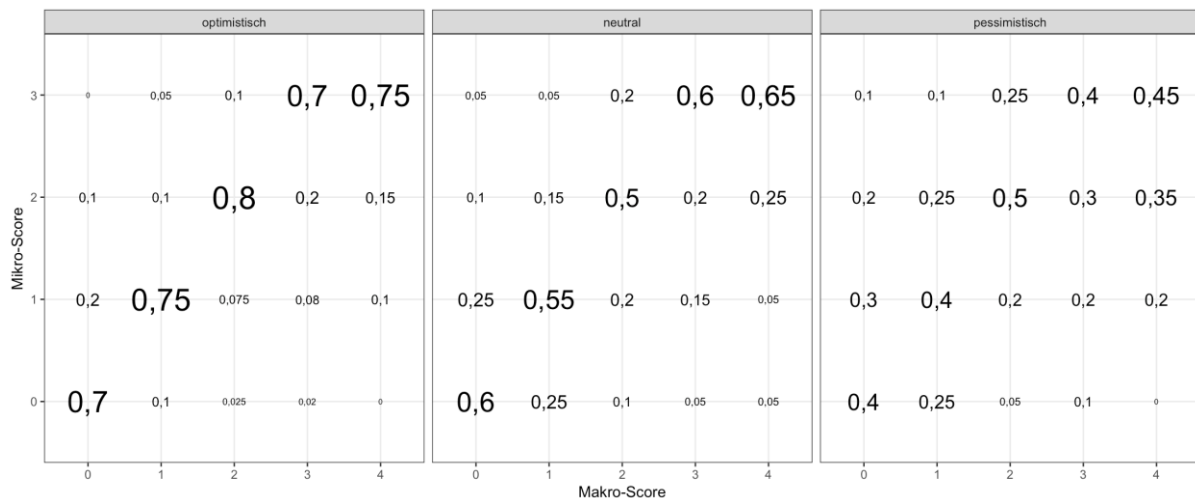
- Implementary Rules of the Lower Saxonian Ministry of Nutrition 2015 *Agriculture and consumer protection for §20 (4; 5). Latest Edition: 31 July 2015, from the German Order on the Protection of Animals and the Keeping of Production Animals (2006)*. [http://www.nds-voris.de/jportal/portal/t/ji3/page/bsvorisprod.psmlijsessionid=8F93DE217C1A385FC1EBBE0F60477.jp27?pid=Dokumentanzeige&show-doccase=I&js\\_peid=Trefferliste&documentnumber=I&numero-fresults=I&fromdoctodoc=yes&doc.id=VVND-VVND000035929#focuspoint](http://www.nds-voris.de/jportal/portal/t/ji3/page/bsvorisprod.psmlijsessionid=8F93DE217C1A385FC1EBBE0F60477.jp27?pid=Dokumentanzeige&show-doccase=I&js_peid=Trefferliste&documentnumber=I&numero-fresults=I&fromdoctodoc=yes&doc.id=VVND-VVND000035929#focuspoint)
- Landis JR and Koch G 1977 The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174. <https://doi.org/10.2307/2529310>
- Louton H, Erhard M and W Öhr A 2018 Acquisition of animal-based welfare measures at slaughter of poultry. *Fleischwirtschaft* 11: 94-98
- Martland MF 1985 Ulcerative dermatitis in broiler chickens: The effects of wet litter. *Avian Pathology* 14: 353-364. <https://doi.org/10.1080/03079458508436237>
- Martrenchar A, Boilletot E, Huonnic D and Pol F 2002 Risk factors for foot-pad dermatitis in chicken and turkey broilers in France. *Preventive Veterinary Medicine* 52: 213-226. [https://doi.org/10.1016/S0167-5877\(01\)00259-8](https://doi.org/10.1016/S0167-5877(01)00259-8)
- Mayne RK, Hocking PM and Else RW 2006 Foot pad dermatitis develops at an early age in commercial turkeys. *British Poultry Science* 47: 36-42. <https://doi.org/10.1080/00071660500475392>
- McKeegan D 2010 *Foot pad dermatitis and hock burn in broilers Risk factors, aetiology and welfare consequences* Research Project Final Report, Faculty of Veterinary Medicine, University of Glasgow, Scotland, UK. <http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=14546>
- Meluzzi A, Fabbri C, Folegatti E and Sirri F 2008 Survey of chicken rearing conditions in Italy: Effects of litter quality and stocking density on productivity, foot dermatitis and carcass injuries. *British Poultry Science* 49: 257-264. <https://doi.org/10.1080/00071660802094156>
- Michel V, Prampart E, Mirabito L, Allain V, Arnould C, Huonnic D, Le Bouquin S and Albaric O 2012 Histologically validated footpad dermatitis scoring system for use in chicken processing plants. *British Poultry Science* 53: 275-281. <https://doi.org/10.1080/00071668.2012.695336>
- R Core Team 2018 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria. <https://www.R-project.org/>
- Saraiva S, Saraiva C and Stilwell G 2016 Feather conditions and clinical scores as indicators of broilers welfare at the slaughterhouse. *Research in Veterinary Science* 107: 75-79. <https://doi.org/10.1016/j.rvsc.2016.05.005>
- Shepherd EM and Fairchild BD 2010 Foot-pad dermatitis in poultry. *Poultry Science* 89: 2043-2051. <https://doi.org/10.3382/ps.2010-00770>
- Spindler B and Hartung J 2011 Prevalence of pododermatitis in broiler chicken kept according to Directive 2007/43/EC Stocking Densities. *XVth ISAH Congress on Animal Hygiene International Society for Animal Hygiene (ISAH)* pp 39-42. 3-7 July, Vienna, Austria
- Weber W yneken C, Sinclair A, Veldkamp T, Vinco LJ and Hocking PM 2015 Foot-pad dermatitis and pain assessment in turkey poults using analgesia and objective gait analysis. *British Poultry Science* 56: 522-530. <https://doi.org/10.1080/00071668.2015.1077203>
- Welfare Quality® 2009 *Welfare Quality® Assessment Protocol for Poultry (Broilers, Laying Hens)*. Welfare Quality® Consortium: Lelystad, Netherlands
- Wideman Jr RF 2016 Bacterial chondronecrosis with osteomyelitis and lameness in broilers: A review. *Poultry Science* 95: 325-344. <https://doi.org/10.3382/ps/pev320>
- Wilcox CS, Patterson J and Cheng HW 2009 Use of thermography to screen for subclinical bumblefoot in poultry. *Poultry Science* 88: 1176-1180. <https://doi.org/10.3382/ps.2008-00446>



## V. ERWEITERTE ERGEBNISSE

### 1. Ergebnisse der Stichprobenberechnung

Zur Berechnung der Stichprobengröße wurden Datensätze generiert, da zu diesem Zeitpunkt noch keine Daten vorlagen. Die Elemente der so erstellten Zuordnungsmatrizen sind die bedingten Wahrscheinlichkeiten für das Auftreten der vier Mikroscore-Ausprägungen in Bezug auf einer der fünf Makroscore-Ausprägungen. Zu einem gegebenen Makroscore (X-Achse) ist die Wahrscheinlichkeit für das Auftreten aller Mikroscore-Ausprägungen (Y-Achse) gegeben. Zudem wurden drei hypothetische Szenarien (optimistisch/ neutral/ pessimistisch) erstellt. Durch diese generierten Datensätze wird die Güte des Verfahrens bei einem speziellen Stichprobenumfang abgeschätzt. Die Zielgröße Mikroscore ist kategorial und ordinal skaliert, jedoch variieren die Wahrscheinlichkeiten entlang der Makro-Score-Kategorien nicht um den gleichen Faktor (keine proportionale „Odds Annahme“) daher wurde ein multinomiales logistisches Regressionsmodell für die Schätzung der Zuordnungsmatrix verwendet. Das Ergebnis der Analyse ist in Abbildung 3 dargestellt.



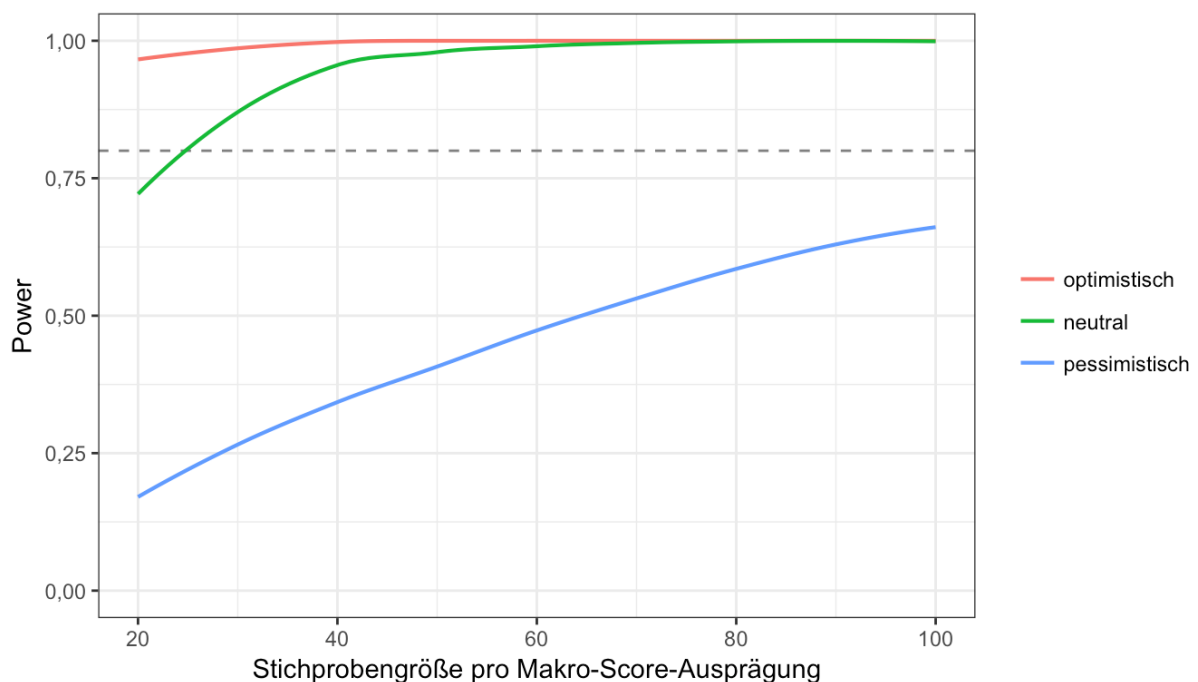
**Abbildung 3:** Zuordnungsmatrizen des Mikroscores zugeordnet dem Makroscore mit optimistischen / neutralen / pessimistischen Szenario.

Zum Beispiel ist die Wahrscheinlichkeit bei dem Mikroscore eine 1 zu erhalten, wenn bei dem Makroscore ebenfalls eine 1 beobachtet wurde im neutralen Szenario 0,55. Dem gegenüber stehen die Mikroscore-Ausprägungen 0, 2 und 3 mit den Wahrscheinlichkeiten 0,25; 0,15 und 0,05. Auf gleiche Art werden die Werte zu den anderen Makroscore-Ausprägungen spaltenweise gelesen. Wie aus den drei Grafiken hervorgeht, wurde die Zuordnungen so



erstellt, dass höhere Makroscore-Werte auch eher mit höheren Mikroscore-Werten verknüpft sind, um so die zu erwartenden Gegebenheiten widerspiegeln zu können.

Der Unterschied zwischen den Grafiken ist der Absolutbetrag der Wahrscheinlichkeiten. Bei allen drei Szenarien lässt sich eine höchste Wahrscheinlichkeit erkennen, jedoch lassen sich nur beim optimistischen Szenario die verschiedene Score-Ausprägungen gut differenzieren, bei der neutralen und pessimistischen Schätzung sind die Wahrscheinlichkeitsunterschiede deutlich geringer, sodass eine eindeutige Zuordnung erschwert wird. Die daraus resultierenden Powerwerte sind in Abbildung 4 dargestellt.



**Abbildung 4:** Schätzung der Power in Bezug auf die Stichprobengröße pro Makroscore-Ausprägung mit optimistischen / neutralen / pessimistischen Szenario.

Der als akzeptabel angesehene Power-Wert (die Wahrscheinlichkeit, dass die Score-Ausprägungen mit den höchsten Wahrscheinlichkeiten korrekt erkannt werden) von 0,8 wird beim optimistischen Szenario bereits erreicht, wenn 20 Untersuchungen pro Makroscore Kategorie durchgeführt werden. Für den neutralen Fall genügen zwischen 20 und 25 Beobachtungen, im pessimistischen Fall werden selbst mit 100 Untersuchungen pro Kategorie keine akzeptablen Power-Werte erreicht.

## 2. Vergleich Kerascore mit der manuell gemessenen relativen Flächenveränderung

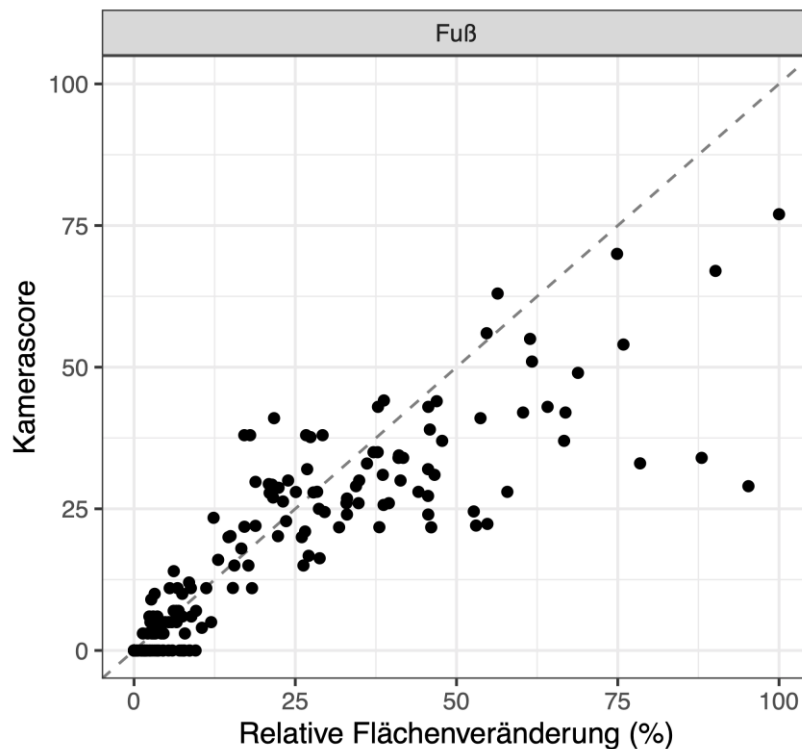
Um die relative Flächenveränderung der Fläche der Läsion ( $A_L$ ) in Bezug auf die Fläche des gesamten Fußballens ( $A_F$ ) mit dem Kerascore vergleichen zu können, müssen die Flächen ( $A_L$  und  $A_F$ ) durch Kreise approximiert werden. Dies ist notwendig, da sowohl für die Läsionen als auch für den Fußballen die maximalen Durchmesser bestimmt wurden ( $d_L$  und  $d_F$ ).

$$\frac{A_L}{A_F} \cdot 100 = \frac{2\pi \left(\frac{d_L}{2}\right)^2}{2\pi \left(\frac{d_F}{2}\right)^2} \cdot 100 = \left(\frac{d_L}{d_F}\right)^2 \cdot 100$$

Dies kann durch Wegkürzen der Konstanten vereinfacht werden:

$$\left(\frac{d_L}{d_F}\right)^2 \cdot 100$$

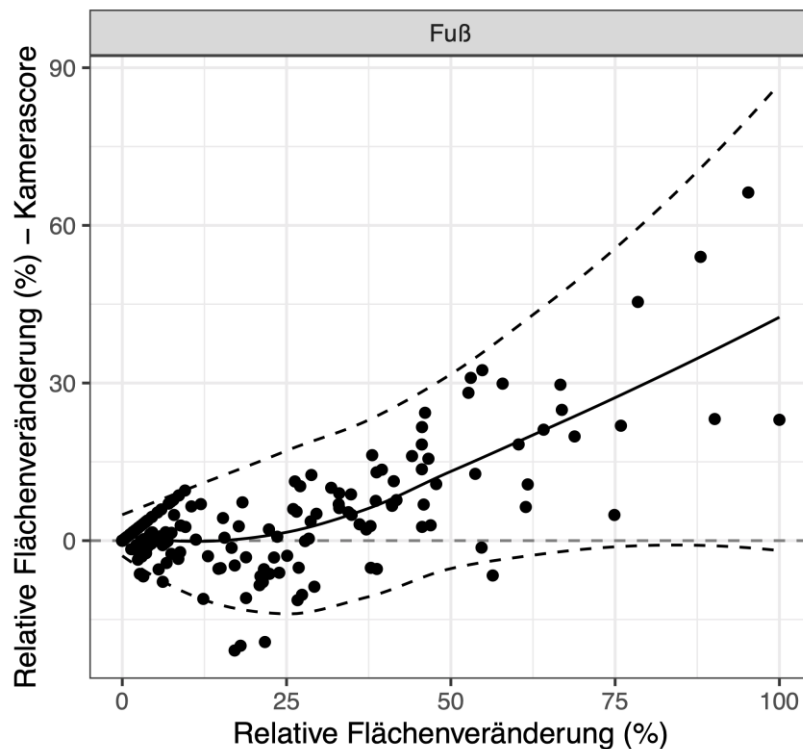
Da sowohl der Kerascore als auch die relative Flächenveränderung metrische Merkmale sind, ist die Analyse der Daten komplex. Der Vergleich dieser beiden metrischen Merkmale wird in zwei Schritten durchgeführt: Zunächst erfolgt eine Ermittlung eines Bias, das bedeutet es wird überprüft, ob eine systematische Abweichung der beiden Scores zueinander vorhanden ist (Abbildung 5).



**Abbildung 5:** Kamerascore in Bezug auf die manuell gemessene, relative Flächenveränderung der Foot-Pad Dermatitis.

Auf der X-Achse ist die manuelle Messung abgetragen, auf der Y-Achse der Kamerascore dargestellt. Die einzelnen Punkte stellen die Differenzen zwischen den Scores dar, wären alle Punkte auf der Hauptdiagonalen (gestrichelten Linie) wäre die beiden Scores identisch, sie würden sozusagen zu 100% gleich gut FPD erfassen. Wie in der Abbildung 5 dargestellt, weicht die Punktwolke vor allem bei hohen relativen Flächenveränderungen in Richtung der X-Achse ab, da fast alle Punkte unter der Diagonalen liegen, was einen Bias (systematische Abweichung) darstellt.

Dies bedeutet zudem, dass der Kamerascore in fast allen Fällen geringere Werte aufweist als bei der manuellen Beurteilung gemessen wurde, dies kommt durch die Verwendung des maximalen Durchmessers bei der Kreisberechnung der relativen Flächenveränderung zustande. Der Mittelwert des Fehlers dieser beiden Methoden beträgt 3,67% und die Standardabweichung 10,64% mit einem Minimalwert von - 20,9% und einem Maximalwert von 66,26%. Um die Grenzen der Übereinstimmung betrachten zu können, wird auf Grundlage der vorliegenden Daten ein Bland-Altman-Plot erstellt (Abbildung 6).



**Abbildung 6 :** Bland-Altman Plot: Die Differenz der relativen Flächenveränderung in % und des Kameranotes in Bezug auf die relative Flächenveränderung der Foot-Pad Dermatitis in %.

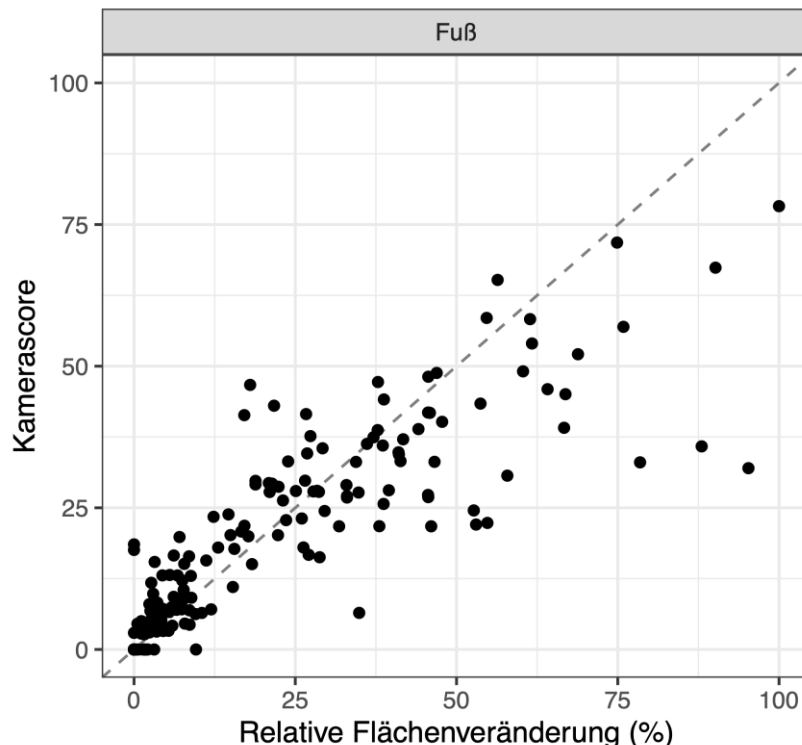
Normalerweise ist auf der Y-Achse nicht der Mittelwert aus beiden Messungen dargestellt, bei Abbildung 6 liegt es daran, dass im vorliegenden Fall die manuelle Messung als Referenz angenommen wird, somit weicht die Darstellung hier von einem üblichen Bland-Altman Plot ab. Eine Gleichheit kann - genau wie bei Abbildung 5 - angenommen werden, wenn die Messungen im Plot annähernd ihre Differenzen um die Nulllinie herum anordnen (Abbildung 6).

Die Streuung der Differenzen als auch die Abweichung der beiden Methoden nimmt mit steigender Flächenveränderung zu. In der Grafik ist dies daran zu erkennen, dass die Punkte (Differenzen) weiter entfernt von der Mittellinie (durchgezogene Linie) liegen.

### 3. Vergleich des verbesserten Kerascores mit der manuell gemessenen relativen Flächenveränderung

Unabhängig von dieser Studie wurde auf Seiten des Kooperationspartners CLK GmbH, welcher die automatisierten Kerasysteme betreut und herstellt, an einer Verbesserung der Aufnahmetechnik gearbeitet. Das Aufnahmeverfahren basiert auf der Suche nach Regionen mit Kontrastunterschieden, welche nach Helligkeit herausgefiltert werden, sodass dunkle Stellen verbleiben. Diese dunklen Areale stellen die Fläche der Läsion dar. Bei dem verbesserten Kerascore wurden nun die Grenzwerte für die maximale Helligkeit erhöht, zudem wurde der Grenzwert für die Erkennung von Kontrastunterschieden verkleinert, sodass auch schon bereits weniger dunkle Bereiche als zuvor erkannt werden.

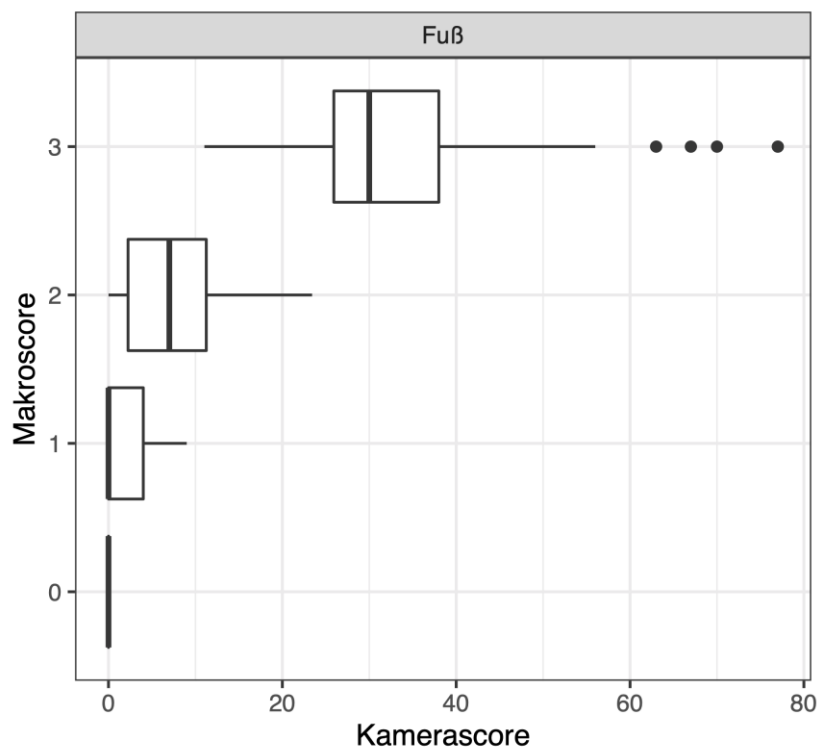
Der Mittelwert des Fehlers dieser beiden Methoden beträgt 1,53% mit einem Minimalwert von 11,11% und einem Maximalwert von -28,70%, sowie einer Standardabweichung von 63,26%. Dies ist in der Abbildung 7 dargestellt. Auf der X-Achse ist die manuelle Messung der relativen Flächenveränderung abgebildet, auf der Y-Achse ist der Kerascore abgetragen. Auch hier gilt analog die Abbildung 5 und 6: Würden beide Methoden in etwa das gleiche messen, so würden die Punkte annähernd symmetrisch um die eingezeichnete Hauptdiagonale streuen. Die Punktwolke liegt nun näher an der der Hauptdiagonale.



**Abbildung 7:** Verbesserter Kerascore in Bezug die manuell gemessene, relative Flächenveränderung.

#### 4. Relation zwischen Kameranumber und Makroscore

Der Makroscore 3 und 4 unterscheidet sich lediglich darin, dass bei Makroscore 4 eine oder mehrere tiefe Läsionen an den Zehen vorliegt, da das Kameranumber die Zehen nicht miterfasst, wurde in den nachfolgenden Betrachtungen Makroscore 3 und 4 zusammengefasst und in den folgenden Abschnitten gemeinsam als Makroscore 3 betrachtet. Wie in der Abbildung 8 dargestellt, zeigt die deskriptive Darstellung die Verteilung des Kameranumber in % für die jeweiligen Makroscore-Kategorien. Ein starker Anstieg des Kameranumber mit steigenden Makroscore-Kategorien ist sichtbar, obwohl keine direkte Proportionalität feststellbar ist.



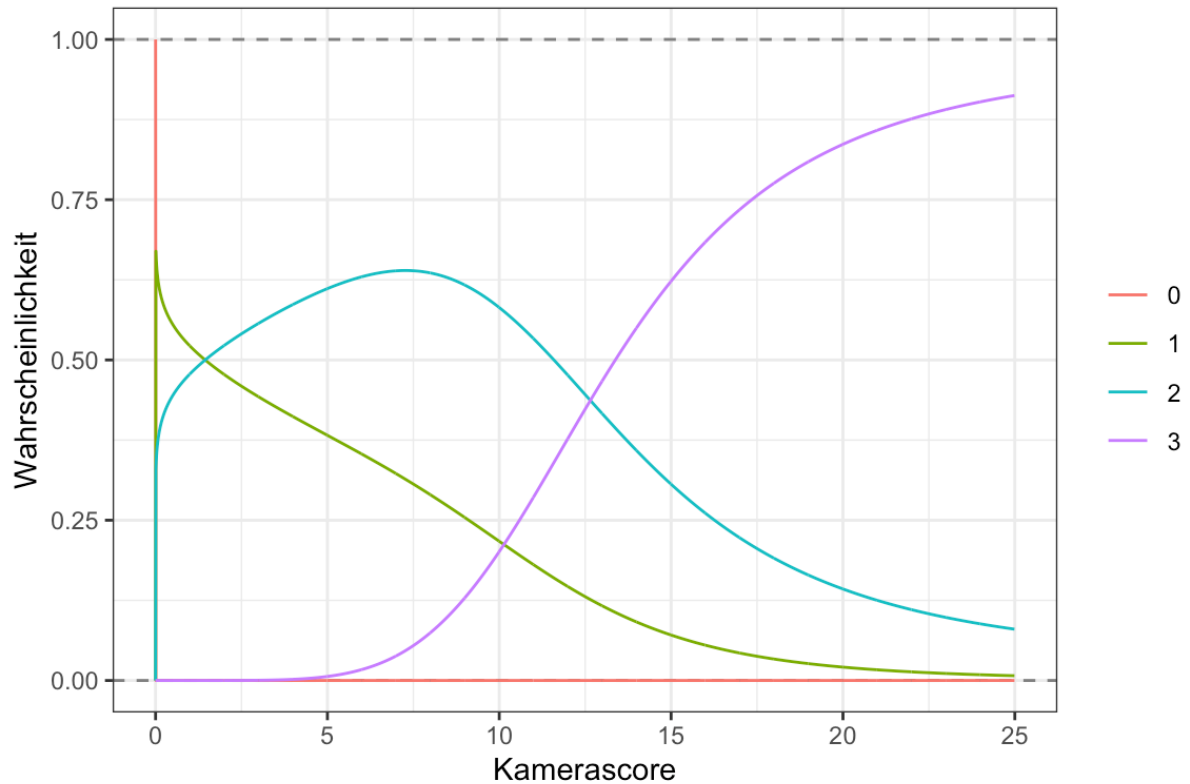
**Abbildung 8:** Deskriptive Darstellung des Makroscores und Kameranumber im Vergleich (Makroscore 3 zusammengefasst, vorher: Makroscore 3 und 4).

Für die Füße erhalten wir die in der Tabelle 12 dargestellte numerische Zusammenfassung des Kameranumber entlang der Makroscore-Kategorien (Makroscore 3 und 4 wieder zusammengefasst). Hier kann man erkennen, dass alle Füße mit einem Makroscore von 0 auch einen Kamerawert von 0 aufweisen. Der Wert für „Kameranumber = 0“ sinkt kontinuierlich mit höheren Makroscores, so ist er bei Makroscore 1 bei 55% und bei Makroscore 2 bei 25%. Bei Makroscore 3 und 4 liegt kein einziger Kamerawert mit 0 vor (0%). Interessant ist zudem, dass die Mittelwerte (Mean) recht deutlich überproportional von 0, über 2,0 und 8,0, auf 33,4 ansteigen.

**Tabelle 12:** Verteilung des Kerascores entlang der Makroscore-Kategorien (n = 200, Mean= Mittelwert, SD =Standardabweichung, Min = Minimalwert, Max= Maximalwert).

Score	N nach Makroscore	Kerascore = 0	Mean	SD	Min	Max
0	40	100%	0,0	0,0	0	0
1	40	55%	2,0	2,5	0	9,0
2	40	25%	8,0	6,8	0	23,4
3 und 4	80	0%	33,4	12,3	11	77,0

Analog wie im Vergleich Mikroscore und Makroscore (in der Publikation Piller et al., 2020) wurde nun durch ein multinomiales logistisches Regressionsmodell mit dem Makroscore als Zielgröße und dem Kerascore als Einflussgröße eine Schätzung der bedingten Wahrscheinlichkeiten für die Makroscore-Kategorien bezogen auf die Kerascore-Werte durchgeführt. Auf Grundlage der in Tabelle 12 dargestellten Verteilung des Kerascores auf die Makroscore-Kategorien lassen sich für die Verteilung Wahrscheinlichkeiten berechnen, welche in der Abbildung 9 veranschaulicht sind.



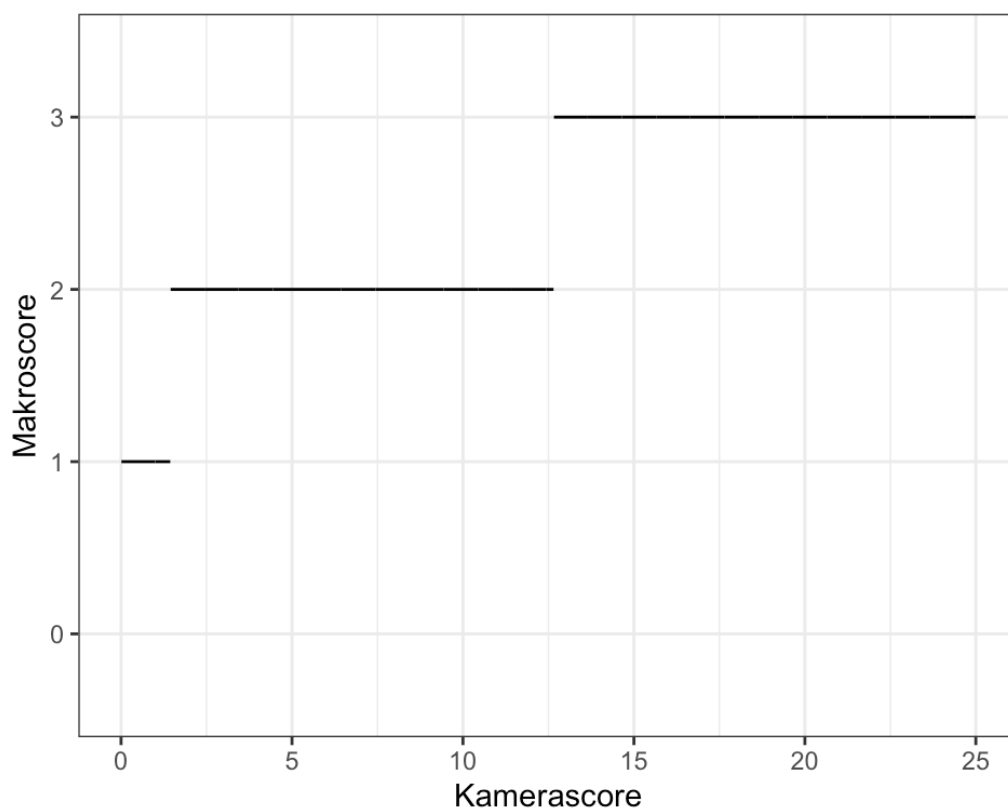
**Abbildung 9:** Wahrscheinlichkeitsschätzung für die Verteilung des Makroscores entlang des Kamerascoringes (n = 200).

In Abbildung 9 ist auf der X-Achse der Kamerascoring in % abgebildet und die Wahrscheinlichkeiten von 0 bis 1 auf der Y-Achse. Jede Makroscore-Verteilung ist gut sichtbar und klar abtrennbar von den anderen Scores. Jeder Makroscore hat einen eigenen Abschnitt des Kamerascoringes bei dem er die höchste Wahrscheinlichkeit aufweist, jedoch ist anzumerken, dass Score 0 und 1 jeweils nur einen sehr kleinen Bereich mit der höchsten Wahrscheinlichkeit abdecken.



## 5. Cut-Off-Werte des Kerascores für den Makroscore

In der Abbildung 10 werden die höchsten Wahrscheinlichkeiten für Makroscore-Kategorien in Bezug auf den Kerascore präsentiert: Auf der X-Achse wurde der Kerascore in % und auf der Y-Achse der Makroscore mit den jeweiligen Kategorien dargestellt. Der Makroscore 0 weist zu 100% auch einen Kerascore von 0 auf, ansteigend dazu weist der Makroscore 1 bei einem Kerascore von 55%, der Makroscore 2 bei einem Kerascore von 25% und der zusammengefasste Makroscore 3 (Makroscore 3 und 4) bei einem Kerascore von 0% die höchste Wahrscheinlichkeit auf. Die Bestimmung der optimalen Cut-Offs führt über die Berechnung der bedingten Wahrscheinlichkeiten für die Makroscore Kategorien 0, 1, 2, und 3 in Bezug auf den Kerascore:



**Abbildung 10:** Höchste Wahrscheinlichkeiten für die Makroscores in Bezug auf den Kerascore (Makroscore 3 und 4 als 3 zusammengefasst).

Die hierdurch abgeleiteten Cut-Offs können der Tabelle 13 entnommen werden. Die so erhaltenen Cut-Off-Werte zeigen, dass dem Makroscore 0 nur Füße mit einem Kerascore von 0 zugeordnet werden. Bereits ab 0,01 Kerascore bis zu einem Kerascore von 1,44 werden die von der Kamera erfassten Läsionen dem Makroscore 1 zugeordnet. Zu bemerken ist hier, dass der Bereich, in dem der Makroscore 1 die höchste Wahrscheinlichkeit aufweist, im Verhältnis zu den folgenden Scores recht klein erscheint, da Makroscore 2 bereits über 10

Kamerascore-Prozentpunkte einnimmt. Jedoch weist bereits ab einen Kamerawert von 12,66 der zusammengefasste Makroscore 3 (Makroscore 3 und 4) die höchsten Wahrscheinlichkeiten auf.

**Tabelle 13:** Cut-Off-Werte für die Verteilung der Makroscores entlang des Kamerascors (n = 200, Min = Minimalwert, Max = Maximalwert).

Makroscore	Min	Max
0	0,00	0,00
1	0,01	1,44
2	1,45	12,65
3 und 4	12,66	Inf

Die Performanzenwerte (Tabelle 14) für Makroscore 0 sowie 3 und 4 zeigen Werte für die Sensitivität für die Diagnostik der jeweiligen Makroscore-Kategorien -bei dem Einsatz von den genannten Cut-Off Werten des Kamerascors- von 0,99 oder höher. Für Makroscore 2 liegt die Sensitivität für das Zuordnen des Makroscore-Wertes zum Kamerascor bei gegebenen Cut-Off-Werten bei 0,52 und bei Makroscore 1 ist sie mit 0,00 mit deutlichem Abstand am geringsten. Die Spezifität für die Diagnostik der jeweiligen Makroscore-Kategorien bei dem Einsatz von den genannten Cut-Off Werten des Kamerascors ist bei allen Makroscore-Kategorien bei oder über 0,80.

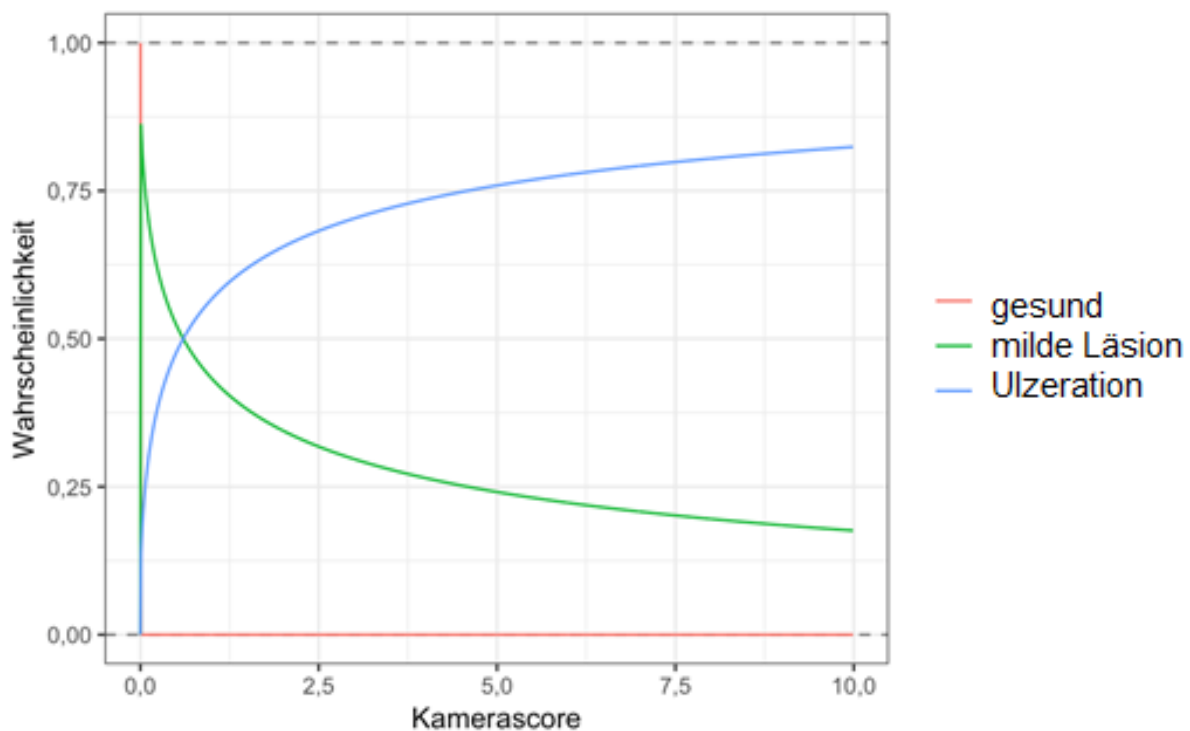
**Tabelle 14:** Für den Makroscore in Bezug auf die Cut-Off-Werte (Makroscore 3 und 4 zusammengefasst, n = 200, KS = Kerascore, Acc = Genauigkeit, Sens = Sensitivität, Spec = Spezifität, PPV = positiver prädikativer Wert, NPV: Negativer prädikativer Wert, value= Wert der Prognosegüte, Performance = Prognosegüte).

Vorhersage					Performance			
	KS >0,01	0,01 <= KS < 1,45	1,45 <= KS < 12,66	KS >= 12,66	Accuracy = 0,70			
Makroscore	0	1	2	3&4	Sens	Spec	PPV	NPV
0	40	0	0	0	1,00	0,80	0,56	1,00
1	22	0	18	0	0,00	1,00	NaN	0,80
2	10	0	21	9	0,52	0,88	0,52	0,88
3 & 4	0	0	1	79	0,99	0,92	0,90	0,99

## 6. Relation zwischen Kameranote und Mikroscore

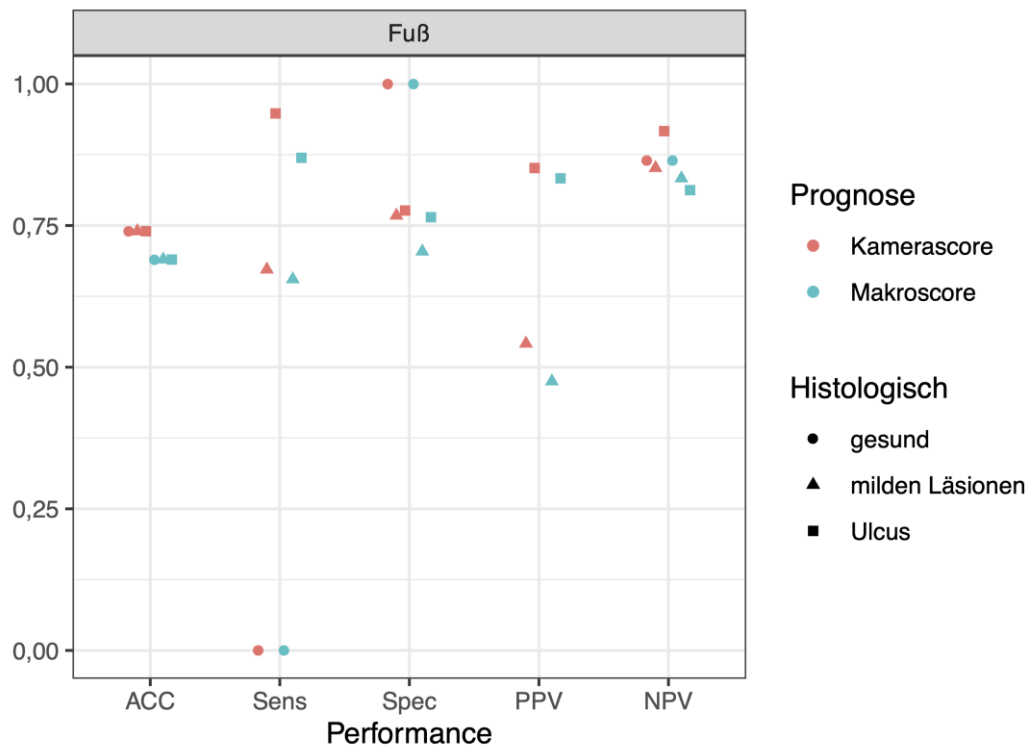
Der Mikroscore wurde wie folgt zusammengefasst: Mikroscore 0 sind histologisch gesunde Füße, Mikroscore 1 und 2 werden als milde Läsionen zusammengefasst und Mikroscore 3 sind histologisch betrachtet Ulzerationen (Tabelle 10). Vergleicht man die Ergebnisse der histopathologischen Untersuchung mit den Kamerawerten, so ergibt sich, dass histologisch gesunde Füße immer mit einem Kameranote von 0,0 bewertet werden. Erosionen oder andere milde Läsionen werden überwiegend auch mit einem Kameranote 0,0 bewertet. Erst Ulcera werden überwiegend mit Kameranoten ungleich 0,0 bewertet.

Aufgrund dieser Zuordnung ist eine Schätzung für die bedingten Wahrscheinlichkeiten der drei Kategorien (gesund, milde Läsionen und Ulzeration) für den Kameranote möglich (Abbildung 11).



**Abbildung 11:** Schätzung bedingten Wahrscheinlichkeiten der histologischen Kategorien (gesund, milde Läsionen oder Ulzerationen) in Abhängigkeit des jeweiligen Kameranotes.

Histologisch gesunde Füße werden nur dem Kamerawert 0,0 zugeordnet. Milde Läsionen zeigen nur in einem kleinen Bereich die höchste Wahrscheinlichkeit, danach überwiegen die Ulzerationen (Abbildung 11). Hier ist es möglich einen Vergleich mit den Performancewerten der Makroscore-Kategorien zu ziehen (Abbildung 12).



**Abbildung 12:** Prognosegüte (Performance) von Kerascore und Makroscore verglichen mit den histopathologischen Ergebnissen (n = 200, ACC = Genauigkeit, Sens = Sensitivität, Spec = Spezifität, PPV = positiver prädikativer Wert, NPV: Negativer prädikativer Wert, value= Wert der Prognosegüte, Performance = Prognosegüte).

Auf der X-Achse sind die verschiedenen Performancewerte gegeben. Eine genaue Erklärung der Performancebegriffe ist in Tabelle I im Anhang zu finden. Die y-Achse zeigt den Wert der Prognosegüte an (value). Beide Erfassungsmethoden (Kamera- und Makroscore) weisen eine ähnliche Prognosegüte in Bezug auf die histopathologischen Untersuchungen auf. Vor allem bei histologisch gesunden Füßen, aber auch bei milden Läsionen und Ulzerationen ist die Übereinstimmung annähernd gleich.

## VI. ERWEITERTE DISKUSSION

### 1. Wiedergabe der manuell gemessen relativen Flächenveränderung durch den Kameroncore

Bei einem Vergleich der manuell gemessen, relativen Flächenveränderung und der Beurteilung mit dem Kameroncore lassen sich vor allem mit steigender Flächenveränderung der Läsionen am Fußballen größere Unterschiede in den Differenzen zwischen den Beurteilungsmethoden feststellen. Würden die manuelle Messung der Flächenveränderung und der Kameroncore gleich messen würden sich alle Punkte um die Hauptdiagonale annähernd symmetrisch verteilen und kaum weit davon entfernt streuen. Dies stellt somit eine symmetrische Abweichung dar, auch Bias genannt. Dieser Bias wurde mit einem Mittelwert des Fehlers von 3,67% und einer Standardabweichung von 10,64% bestimmt. Nun gilt es zu diskutieren, ob diese Abweichung und diese Intervallgrenzen als akzeptabel angesehen werden können. Dazu gilt es vorab nochmal zu erwähnen, dass die relative Flächenveränderung manuell gemessen wurde. Die Entfernung der am weitesten entfernten Punkte am Fußballen wurde in Relation zur Entfernung der am weitesten entfernten Punkte der Läsion gesetzt. Weder der Fußballen eines Masthuhns noch die Läsion der FPD sind jedoch exakt kreisförmig, daher handelt es sich nur um eine Schätzung der relativen Flächenveränderung und nicht um eine exakte Messung. Zudem ist durch die Verwendung des maximalen Durchmessers davon auszugehen, dass der errechnete Flächeninhalt des Kreises größer ausfällt als die tatsächliche Fläche. Es ist mit herkömmlichen Mitteln visuell nicht möglich die genaue Fußballenfläche sowie Läsionsfläche am Schlachthof vor Ort genau zu bestimmen, da die Flächen immer asymmetrisch und aufgrund ihrer Natur nie exakt geometrische Formen annehmen.

Daher erscheint die angewendete Methode, über die längste Stelle die relative Flächenveränderung zu schätzen als plausibel und praktikabel, jedoch stellt die Kamera die tatsächliche Fläche der Läsion anhand des erstellten Fotos dar und schätzt nicht die Fläche über den Umweg der Kreisberechnung.

Hier sollte diskutiert werden, ob diese Grenzbereiche mit den theoretischen Grenzwerten übereinstimmen. Es gibt bisher keine Literatur zu Grenzwerten, weder für den Makroscore noch dem Mikroscore, für das untersuchte Kameroncore sowie für Kameroncores die FPD erfassen im Allgemeinen. Daher gestaltet es sich schwierig im Vorhinein einen theoretischen Grenzwert festzulegen. Mit diesem Hintergrund kann man bei den oben berechneten Werten nur vermuten, dass die berechneten „Limits of agreement“ ausreichend sind und die relative Flächenveränderung durch die Kreisschätzung über Durchmesser wiedergegeben werden kann.

Neben dem Stand der Technik muss auch der wissenschaftliche Stand berücksichtigt werden und da beide einem stetigen Wandel unterliegen und auch die Grenzwerte von Kamerasystem zu Kamerasystem unterschiedlich sein können, lässt sich pauschal keine Aussage treffen. Wichtig ist technisch zu definieren, bis zu welcher Größe kleine Veränderungen noch sicher erkannt werden können. Wissenschaftlich gesehen ist es wichtig zu definieren, ab welcher Flächenveränderung eine tierschutzrelevante Veränderung vorliegt. Auf Grundlage dieser Daten sollten dann theoretische Grenzwerte berechnet werden und versucht werden auf Kamerasysteme zu übertragen.

## **2. Wiedergabe der manuell gemessen relativen Flächenveränderung durch den überarbeiteten Kamerascore**

Zudem wurden von der CLK GmbH die Softwareeinstellungen angepasst, um eine höhere Sensitivität bei der Kontrasterkennung zu generieren. Dieses Verfahren könnte zwar ein erhöhtes Risiko für das fehlerhafte Erkennen von Schatten darstellen, jedoch lässt sich eine Verbesserung des Kamerascores in Bezug auf die relative Flächenveränderung feststellen. Der Mittelwert des Fehlers betrug vorher 3,67% und nach Korrektur 1,53%, bei einer geringgradigen Erhöhung der Standardabweichung von 10,64% auf 11,11%.

Es gilt anzumerken, dass mit diese berechneten Werten nur eine Aussage über das angewendete Kamerasystem an genau diesem Standort zulässig ist, eine generelle Übertragung auf andere Systeme oder Softwaremodelle ist nicht oder nur eingeschränkt möglich. Zudem ist die Schätzung der relativen Fläche der Fußballenfläche und der Läsionsfläche über die geschätzte Kreisberechnung durchaus anwenderbar, aber –wie unter VI 1 bereits erwähnt- ist eine direkte Erfassung der Fläche durch die Kamera logischerweise zu bevorzugen. Umso wichtiger ist es, das Kamerasystem regelmäßig zu warten und den Erfassungsort gut zu beleuchten (CLK GmbH, 2015). Auch eine hohe Luftfeuchtigkeit, in der Schlachtkette verdrehte Füße oder andere Umgebungseinflüssen sollten bei Validierungen mit in Betracht gezogen werden, um Fehlerquellen zu minimieren.

### 3. Vergleich Makroscore und Kameranote

Ein Anstieg des Kameranotes ist in Bezug auf die Verteilung der einzelnen Makroscore-Kategorien gegeben, das bedeutet höhere Kameranote-Werte gehen mit größeren Makroscores einher. Eine Eins-zu-Eins-Zuordnung ist allerdings nicht möglich, es ist im Gegenteil sogar ein relativ steiler Anstieg zu sehen. Die Bereiche der höchsten Wahrscheinlichkeiten für den Makroscore 0 und den Makroscore 1 sind sehr klein. Für den Makroscore 0 liegt die höchste Wahrscheinlichkeit sogar nur bei dem Kameranote 0. Auch für den Makroscore 2 ist der Bereich der höchsten Wahrscheinlichkeit des Kameranotes eher gering im Vergleich mit dem Bereich für den Makroscore 3 und 4. Dies bedeutet, dass ab einem Kameranote von 12,66 % die Wahrscheinlichkeit für eine Läsion mit einem Makroscore von 3 oder 4 am höchsten ist (Tiefe Läsionen über 1,0 cm oder Läsion am Fußballen und eine oder mehrere tiefe Läsionen an den Zehen). Betrachtet man die Performance-Werte, so lässt sich mit einer Sensitivität von 1,00 und einer Spezifität von 0,80 der Makroscore 0 bei einem Kamerawert von  $< 0,01$  diagnostizieren. Generell ist es schwierig hier vergleichbare Literatur zu finden, da diese Kamerawerte - wie bereits erwähnt- nur auf das vorliegende Kamerasystem anzuwenden sind und nur wenig Literatur vorhanden ist. Vanderhasselt et al. (2013) untersuchten ein anderes Kamerasystem und kamen zum Schluss, dass wenn offensichtliche Fehleinstufungen durch das System nicht berücksichtigt wurden, die Ergebnisse, vergleichbar mit denen von Experten erhobenen Befunde sind.

### 4. Vergleich des Kameranotes mit dem Makroscore in Bezug auf die histologischen Ergebnisse

Sowohl der Kameranote als auch der Makroscore können die histopathologischen Gegebenheiten mit Einschränkungen widerspiegeln. Wie in der Veröffentlichung Piller et al. (2020) erwähnt, sind die Performance-Werte für den Makroscore im Bereich der milden Läsionen nur mäßig (Sensitivität 0,41). Hier wurde versucht eine Erhöhung der Genauigkeit durch Verschiebung der Score-Grenzen zu erreichen. Generell kann man drauf schließen, dass wie in dieser Studie (Piller et al. 2020) gezeigt wurde, bei relativ kleinen Läsionen mit einem Makroscore 2 (Tiefe Läsionen über 0,5 cm bis zu 1,0 cm oder oberflächliche Läsion über 0,5 cm vgl. Tabelle 10) die höchste Wahrscheinlichkeit für Ulzerationen vorliegen. Das Erfassen von gesunden Füßen sowie von Füßen mit Ulzerationen ist sowohl durch die Kamera als auch visuell, mit ähnlichen Performancewerten möglich.



## 5. Vor- und Nachteile einer kamerabasierten Erfassung am Schlachthof

Es gibt signifikante Unterschiede zwischen der Erhebung von Experten und von Schlachthofmitarbeitern bei der Erfassung der FPD (Martrenchar et al., 2002). Oliveira et al. (2017) fanden heraus, dass Schlachthofmitarbeiter dazu tendieren das Auftreten und den Schweregrad von FPD zu unterschätzen. In der vorliegenden Studie ist eine gute Zuordnung des von einem Experten beurteilten Makroscore zu dem Kamerascore (mit Ausnahme für milde Läsionen) festgestellt worden. Dagegen stimmten in einer anderen Studie automatisch erfasste Scores nur gering mit den von Experten im Stall (54%) und am Schlachthof (59%) überein (Vanderhasselt et al., 2013). De Jong et al. (2015) zählt die Vorteile für standardisierte Messungen von Tierwohlindikatoren am Schlachthof auf: einheitliche Umgebung, bessere Sichtbarkeit der Broiler sowie das Fehlen von Federn und Schmutz nach dem Brühen. Zudem kommen keine Fehlerquellen durch den menschlichen Beobachter wie Müdigkeit, Konzentrations-schwankungen oder Subjektivität der Beurteilung zum Tragen (Bergsma et al., 2015). Auch Louton et al. (2020a) empfehlen eine kamerabasierte Erfassung von FPD am laufenden Schlachtband oder zumindest ein Aushängen aus der Kette, sollte die Beurteilung visuell durch einen Beobachter erfolgen.

Es werden jedoch auch die Nachteile der Beurteilung der FPD am Schlachthof benannt: durch hohe Schlachtgeschwindigkeiten läuft die Schlachtkette sehr schnell am Beobachter / der Kamera vorbei zudem kann an toten Tieren kein Gangbild oder Verhalten mehr beobachten werden (De Jong et al., 2015).

Die letzten beiden Punkte müssen aber unter dem Gesichtspunkt der Prävention betrachtet werden. In erster Linie geht es um Feedback für den Mäster (und womöglich auch Behörden), so können diese für zukünftige Mastdurchgänge Präventionsmaßnahmen einleiten.

Die Übereinstimmung von Beurteilungen durch Experten im Stall und im Schlachthof war nur gering (Ben Sassi et al., 2016). Auch Louton et al. (2020a) kam zu dem Ergebnis, dass visuelle Untersuchungen am Schlachtband zu größeren Abweichungen (bis zu 20%) im Vergleich mit den Untersuchungen im Stall, der visuellen Beurteilung von entnommen Fußpaaren am Schlachthof oder einem Kamerasystem führen. Daher kann eine visuelle Beobachtung direkt am Schlachtband durch Beobachter nicht empfohlen werden, hingegen wies das untersuchte Kamerasystem am Schlachthof ähnliche Beurteilungen wie die im Stall erhobenen Befunde auf und wurde daher empfohlen. Hier ist noch zu erwähnen, dass eines der beurteilten Systeme bei Louton et al. (2020a) ebenfalls das Kamerasystem „Chicken Check“ der Firma CLK GmbH, Altenberge, Deutschland war und sich daher die Ergebnisse gut mit den in dieser Studie erlangten Erkenntnissen vergleichen lässt.

## **6. Vergleich der Ergebnisse für Footpad Dermatitis mit Ergebnissen bei Hock Burn**

Bei einer vergleichbaren Studie über Hock Burn (Louton et al., 2020b) wurden ähnliche Ergebnisse festgestellt wie in der Publikation des Ergebnisteils Piller et al. (2020). Louton et al. (2020b) berichten über einen noch stärkeren Anstieg der Makroscore-Kategorien in Bezug auf den Mikroscore bei der Untersuchung der Fersenhöckerläsionen. In beiden Studien wurde zudem ein Anstieg in der histologisch gemessenen Tiefe der Läsionen in Bezug auf den Makroscore gemessen. Zum Vergleich: Makroscore 3 Läsionen am Fußballen wiesen im Durchschnitt eine Tiefe von 1443  $\mu\text{m}$  auf (Piller et al. 2020), Fersenhöcker mit einer Makroscore 3 Läsion hingegen 1019  $\mu\text{m}$  (Louton et al., 2020b). Nichtsdestotrotz könnte man laut Louton et al. (2020b) in Erwägung ziehen, die Tiefe der Läsion in Zusammenhang mit dem Schweregrad der Läsion zu bringen, da sowohl bei der FPD als auch bei Hock Burn ein deutlicher Anstieg der Tiefe zu erkennen ist, auch wenn dieser nicht signifikant war (Piller et al. 2020). Beide Publikationen beschäftigen sich mit vergleichbaren makroskopischen und mikroskopischen Scoringsystemen. Ebenso wie die FPD (Allain et al., 2009) ist Hock Burn als Tierschutzindikator für Masthühner geeignet (Hepworth et al., 2011).

## 7. Ausblick

Aufgrund der Ergebnisse dieser Studie kann mit hoher Wahrscheinlichkeit das Vorhandensein von histologischen Ulzerationen und histologisch gesunden Füßen durch makroskopische Erfassungsmethoden geschätzt werden, jedoch ist eine 100% sichere Aussage der histologischen Befunde über eine makroskopische Betrachtung (durch Kameras oder Beobachter) nicht möglich. Es ist darüber zu diskutieren ob die in Deutschland gesetzlich in den Ausführungshinweisen der Tierschutz-Nutztierhaltungsverordnung für Masthühner (2011) verankerten histologischen Begriffe „Ulzeration“ und „Erosion“ (letztere wird zu den milden Läsionen gezählt) bei der Beurteilung der FPD berücksichtigt werden sollten oder eher die makroskopische Einteilung. In den Ausführungshinweisen der Tierschutz-Nutztierhaltungsverordnung für Masthühner (2011) sind neben der Nennung histologischer Begriffe auch makroskopische Beispielbilder gegeben. Eine Schmerzhaftigkeit von Ulzerationen scheint laut Literatur vorhanden zu sein, zumal Gentle et al. (1992) feststellte, dass keine großen Unterschiede in der Anatomie, der Physiologie sowie dem Verhalten in Bezug auf Schmerz bei Säugetieren und Vögel vorhanden sind. Jedoch sollte in weiterführenden Untersuchungen genauer erfasst werden, ob Ulzerationen unterschiedlicher Schweregrade unterschiedlich schmerzhaft sind. Es ist zu diskutieren ob eine Ulzeration mit einer Größe von 2 mm ebenso schmerzhaft ist wie eine Ulzeration von 2 cm. Daher scheint eine einheitliche Beurteilung aufgrund der makroskopischen Befunde als geeignet, um FPD darzustellen, zumal eine histologische Auswertung aus Zeit- und Kostengründen flächendeckend nicht anwendbar ist. Eine makroskopische Bewertung scheint mit der Zuhilfenahme von Kamerasystemen praktikabel und zukunftsorientiert sowie objektiv vergleichbar.

## VII. ZUSAMMENFASSUNG

### **Validierung eines makroskopischen und eines kamerabasierten Bewertungssystems der Foot-pad Dermatitis bei Masthühnern durch histopathologische Untersuchungen**

Mit steigenden Schlachtzahlen in Deutschland rückt auch immer mehr die automatisierte Erfassung von Tierwohlindikatoren in den Vordergrund. Durch standardisierte Erfassungsmethoden ist es möglich eine Vielzahl an Tieren, bis hin zu ganzen Herden zu beurteilen. Zudem ermöglichen solche Systeme eine sehr gute Vergleichbarkeit der erhobenen Befunde. In dieser Studie sollte untersucht werden, in wie weit die von einem Kamerasystem erfasste Größe der Flächenveränderung der Foot-pad Dermatitis mit der makroskopischen Bewertung, der histologisch messbaren Tiefe sowie dem Grad der Verletzung (gesund, milde Läsion, Ulcus) zusammenhängt und diese Befunde widerspiegeln kann.

Das Projekt wurde über das Programm zur Innovationsförderung des BMEL über die Bundesanstalt für Ernährung und Landwirtschaft gefördert (FKZ 2817903715). Im Vorfeld der Datenerhebung fand ein Beobachterabgleich in Kooperation mit der Tierärztlichen Hochschule Hannover sowie eine Berechnung des Stichprobenumfanges statt.

Insgesamt wurden jeweils 40 Füße jeder Kategorie eines 5-stufigen makroskopischen Scores (200 insgesamt) von Ross 308 Masthühner im Alter von ca. 40 Tagen durch eine Kamera am laufenden Schlachtband beurteilt. Die Kamera „1.3 MP Farbkamera“ von IDS Imaging, Obersulm erfasst mit Hilfe der Software „Chicken Check“ der CLK GmbH, Altenberge und „Halcon“ von MCTec Software GmbH, München den prozentualen Unterschied zwischen der Läsionsfläche und des gesamten Fußballens.

Anschließend wurden die Läsionen der FPD manuell vermessen, visuell bewertet und histologisch untersucht. Für die visuelle Beurteilung der FPD wurde ein 5-stufiges, modifiziertes Bewertungsschema des Welfare Quality® assessment protocol for Poultry (2009) eingesetzt, für die histologische Untersuchung ein 4-stufiges Schema von Michel et al. (2012). Zusätzlich wurde die Entzündungstiefe bei vorhandener Entzündung gemessen.

Bezüglich der Vergleichbarkeit der makroskopischen Beurteilung mit der mikroskopischen Beurteilung ist ein linearer Anstieg sichtbar, sogleich eine Eins-zu-Eins-Zuordnung nicht möglich ist. Betrachtet man den makroskopischen Score vergleichend gegenüber dem mikroskopischen Score, können gesunde Füße mit deutlicher Mehrheit der Kategorie 0 zugeteilt werden. Das Vorhandensein von Ulzerationen beginnt bei Makroscore 1 und

überwiegt bei Makroscore 2 (58%) und Makroscore 3 (92,5%). Füße mit dem Makroscore 4 weisen ausschließlich Ulzerationen auf (100%).

Bei der gemessenen Entzündungstiefe konnte bei vorhandenen Entzündungen kein signifikanter Unterschied jedoch ein Anstieg mit steigendem Mikroscore festgestellt werden. Die Entzündungen der FPD mit Makroscore 3 (1442,99  $\mu\text{m}$ ) und 4 (1422,17  $\mu\text{m}$ ) waren ähnlich tief.

Bei Fußballen ohne Läsion ist der Kameronascore fast immer 0, bei milden Läsionen weist der Score nur relativ niedrige Performance-Werte auf und bei Ulzerationen sind sehr hohe Performance-Werte zu finden. Die Kamera kann nicht zwischen dem Makroscore 3 und 4 unterscheiden, da dieser über eine tiefe Läsion an den Zehen definiert ist und die Kamera nur den Fußballen, ohne Zehen, erfasst.

Lediglich bei der Bewertung von Füßen mit einem Mikroscore von 1 und 2, weist der Kameronascore, genauso wie die visuelle Beurteilung einige Ungenauigkeiten auf, wobei bei beiden Methoden (visuell durch Beobachter oder durch das Kamerasystem) gesunde Füße und Füße mit Ulzerationen gut diagnostiziert werden können. Die berechnete Prognosegüte zeigt, dass gesunde Füße bei den genannten Cut-Off-Werten bei der makroskopischen Beurteilung (Piller et al., 2020) mit einer Sensitivität von 100% und Ulcera zu 95% richtig zu geordnet werden. Dahingegen können Erosionen nicht diagnostiziert werden (Sensitivität = 0%).

Insbesondere das Vorhandensein eines Ulcus verursacht mit hoher Wahrscheinlichkeit Schmerzen und hat somit eine Tierschutzrelevanz. Die vorliegende Dissertation zeigte, dass die flächendeckende Erfassung von Ulzerationen der FPD mit Hilfe der kamerabasierten Technik ermöglicht werden kann.

## VIII. SUMMARY

### **Validation of a macroscopic and a camera based scoring system for foot-pad dermatitis in broiler chickens by using histopathological findings**

Automated systems for the assessment of animal welfare indicators are getting more important due to the high number of animals slaughtered in Germany. It is nowadays possible to score high numbers of broilers, up to whole flocks, in a short amount of time by using automated scoring systems. In addition, those automated scoring systems generate a high comparability of the diagnosis. The aim of the study was to examine, if the camera based scoring of foot-pad dermatitis lesions can be compared to a macroscopic score for FPD, a measurement of the histological depth of a lesion and to the histopathological grade of severity of the lesion (no lesion, mild lesion, ulcers).

The project was supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme (Grant Number 2817903715). Prior to data collection, the interobserver reliability was tested in cooperation with the University of Veterinary Medicine Hannover as and a calculation of the sample size was performed.

In total, 200 feet (40 of each score) from 39-42 day old Ross 308 broiler chickens were scored by a “1.3 MP” color camera by IDS Imaging, Obersulm, Germany with the software “Chicken Check” by CLK GmbH, Altenberge, Germany and the software “Halcon” by MCTec Software GmbH, Munich, Germany at the slaughter line (measuring the area of the lesion compared to the total plantar foot area). Additionally, the lesions of all feet were measured manually, scored visually and histological slides of the lesions prepared. These slides with hematoxylin-eosin stain were scored according to the scheme by Michel et al. (2012) and the depth of inflammation was measured. In order to reduce bias, each histological scoring was assigned to a single foot, but took place without knowing the macroscopic score.

The results show, that a one to one assignment is not possible but a linear increase is visible, if macroscopically findings are compared to microscopically findings. Feet with no visible lesion are in majority assigned to histological feet without lesions. The presence of ulcer starts with macroscopic score 1, is predominant at macroscopic score 2 (58%, e.g. superficial lesions with a lesion over 0.5 cm diameter) and macroscopic score 3 (92.5 %). Feet with macroscopic score 4 only contained histopathological scored ulcers (100%).

Even if not significant, an increase of the inflammation depth was observed in higher macroscopic scores. The inflammation depths from macroscopic score 3 (1442.99  $\mu\text{m}$ ) and macroscopic score 4 (1422.17  $\mu\text{m}$ ) were similar.

For feet without lesions, the camera based score was nearly always 0. Relative low performance values considering the diagnosis were found for mild lesions and very high ones for ulcers. Macroscopic score 4 is defined as having -additionally to the core lesion on the footpad- a lesion or several ones on the toe. The used camera system could not distinguish between macroscopic score 3 and 4, because the camera system only detects the footpad, not the toes.

Both scoring systems (camera based and visually assessed) showed some limitations in detecting feet with microscopic score 1 and 2, but could diagnose healthy feet and feet with ulcers with good performance values. The sensitivity e.g. for the macroscopic scoring were very high for feet with no lesion (100%) and ulcers (95%), but very poor for mild lesions (0%).

Especially the presence of an ulcer is very likely to be a painful lesion and should be considered as an animal welfare issue. The study proved that these can be diagnosed by cameras at slaughter.

## IX. LITERATURVERZEICHNIS

Algers B, Berg C.(2001) Monitoring animal welfare on commercial broiler farms in Sweden. Acta Agric. Scandinavica Section A. Animal Science, 51: 88-92

Allain V, Mirabito L, Arnould C, Colas M, Le Bouquin S, Lupo C, Michel V. (2009) Skin lesions in broiler chickens measured at the slaughterhouse: relationships between lesions and between their prevalence and rearing factors. British Poultry Science, 50: 407-417

Ask B. (2010) Genetic variation of contact dermatitis in broilers. Poultry Science, 89: 866-875

Ausführungshinweisen der Tierschutznutztierhaltungsverordnung für Masthühner. (2011) Beschlossen von der AG Tierschutz der LAV am 29. November 2011. Tierschutz-Nutztierhaltungsverordnung (TierSchNutzTV) i. d. F. der Bekanntmachung vom 22. August 2006 (BGBl. I S. 2043), die durch die Verordnung vom 1. Oktober 2009 (BGBl. I S. 3223) geändert worden ist, Abschnitt 4, Anforderungen an das Halten von Masthühnern.

Ben Sassi N, Averós X, Estevez I. (2016) Technology and Poultry Welfare. Animals, 6: 62

Bergsma J, Heemskerk W, Pieterse C, Hupkes H. (2015) Improving animal welfare: Automated monitoring of footpad lesions. HSA International Symposium 2015: Recent Advances II, 44

Bilgili S, Alley M, Hess J, Nagaraj M. (2006) Influence of age and sex on footpad quality and yield in broiler chickens reared on low and high density diets. Journal of Applied Poultry Research, 15: 433-441

Bland J M, Altman D G. (1986) Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. The Lancet, 327: 307-310

Bruce D W, McIlroy S G, Goodall E A. (1990) Epidemiology of a contact dermatitis of broilers. Avian Pathology, 19: 532-537

CLK GmbH. (2015) HANDBUCH „ChickenCheck“ Dokumentennummer: 20100120, 48



De Jong I C, Hindle V A , Butterworth A, Engel B, Ferrari P, Gunnink H, Perez Moya T, Tuytens F A M, van Reenen C G. (2016) Simplifying the Welfare Quality® assessment protocol for broiler chicken welfare. *Animal*, 10: 117-127

Hepworth P J, Nefedov A V, Muchnik I B, Morgan K L. (2011) Hock burn: an indicator of broiler flock health. *Veterinary Record*, 168: 303

Ekstrand C, Algers B, Svedberg J. (1997) Rearing conditions and foot-pad dermatitis in Swedish broiler chickens. *Preventive Veterinary Medicine*, 31: 167-174

EU-Generaldirektion Gesundheit und Lebensmittelsicherheit. (2016) Use of slaughterhouse data to monitor welfare of broilers on farm, Paper version: ISBN 978-92-79-53346-4

Gentle M J. (1992) Pain in birds. *Animal Welfare*, 1: 235-247

Greene J A, McCracken R M, Evans R T. (1985) A contact dermatitis of broilers -clinical and pathological findings. *Avian Pathology*, 14: 23-38

Haslam S M, Knowles T G , Brown S N, Wilkins L J, Kestin S C, Warriss P D, Nicol C J. (2007) Factors affecting the prevalence of foot pad dermatitis, hock burn and breast burn in broiler chicken. *British Poultry Science*, 48: 264-275

Heitmann S, Stracke J, Petersen H, Spindler B, Kemper N. (2018) First approach validating a scoring system for foot-pad dermatitis in broiler chickens developed for application in practice. *Preventive Veterinary Medicine*, 154: 63-70

Kyvsgaard N C, Jensen H, Ambrosen T, Toft N. (2013) Temporal changes and risk factors for foot-pad dermatitis in Danish broilers. *Poultry Science*, 92: 26-32

Louton H, Erhard M, Wöhr A. (2018) Acquisition of animal-based welfare measures at slaughter of poultry. *Fleischwirtschaft*, 11: 94-98

Louton H, Erhard M, Wirsch K, Bergmann S, Piller A, Schmidt P, Rauch E. (2020a) Vergleich von vier Beurteilungsmethoden zur Erfassung von Kontaktdermatitiden am Fußballen und

Fersenhöcker von Masthühnern. Berliner und Münchner Tierärztliche Wochenschrift, 133: 1-11

Louton H, Piller A, Bergmann S, Erhard M, Stracke J, Spindler B, Kemper N, Schmidt P, Schade B, Boehm B, Kappe E, Bachmeier J, Schwarzer A. (2020b) Histologically validated scoring system for the assessment of hock burn in broilers. Avian Pathology, DOI: 10.1080/03079457.2020.1712328

McKeegan D. (2010) Foot pad dermatitis and hock burn in broilers: risk factors, aetiology and welfare consequences. Research Project Final Report, Faculty of Veterinary Medicine, University of Glasgow, UK Letzter Aufruf: 23.02.2020 <http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=14546>

Martland M F. (1985) Ulcerative dermatitis in broiler chickens: The effects of wet litter. Avian Pathology, 14: 353-364

Matenche A, Boilletot E, Huonnic D, Pol F. (2002) Risk factors for foot-pad dermatitis in chicken and turkey broilers in France. Preventive Veterinary Medicine, 52: 213-226

Meluzzi A, Fabbri C, Folegatti E, Sirri F. (2008) Survey of chicken rearing conditions in Italy: effects of litter quality and stocking density on productivity, foot dermatitis and carcass injuries. British Poultry Science, 49: 257-264

Michel V, Prampart E, Mirabito L, Allain V, Arnould C, Huonnic D, Le Bouquin S, Albaric O. (2012) Histologically-validated footpad dermatitis scoring system for use in chicken processing plants. British Poultry Science, 53: 275-281

Nickel R, Schummer A, Seiferle E. (1992) Lehrbuch der Anatomie der Haustiere, Band V Anatomie der Vögel, Verlag Paul Parey, 27, ISBN: 3-489-57616-0

Oliveira A R S, Lund V P, Christensen J P, Nielsen L R. (2017). Inter-rater agreement in visual assessment of footpad dermatitis in Danish broiler chickens. British Poultry Science, 58: 224-229

Pschyrembel. (2002) Klinisches Wörterbuch, ISBN 978-31101721332002, Gruyter, 334

Qin S M, Bai W Q, Zhang K Y, Ding X M, Bai S P, Wang J P, Peng H W, Yang Y F, Chen C, Zeng Q F. (2019) Different microbiomes are found in healthy breeder ducks and those with foot pad dermatitis. Poultry Science, 98:6340-6348

R Core Team. (2017) R: A Language and Environment for Statistical Computing. Vienna, Austria: RFoundation for Statistical Computing. letzter Aufruf: 20.02.2020 <https://www.R-project.org>

Runderlass des Niedersächsischen Ministeriums für Ernährung, Landwirtschaft und Verbraucherschutz (2014) Ausführungsbestimmungen zur Umsetzung von §20 Abs 4 und 5 der TierSchNutzV (Haltung von Masthühnern). RdErl. d. ML v. 11. 12. 2014 – 204.1-42503/2-828. Erlassdatum 11.12.2014. VORIS 78530. Internetzugriff am 10.03.2020: <http://www.voris.niedersachsen.de/jportal/?quelle=jlink&psml=bsvorisprod.psml&feed=bsvors-vv&docid=VVND-VVND000041583>

Saraiva S, Saraiva C, Stilwell G. (2016) Feather conditions and clinical scores as indicators of broilers welfare at the slaughterhouse. Research in Veterinary Science, 107: 75-79

Shepherd E M, Fairchild B D. (2010) Footpad dermatitis in poultry. Poultry Science, 89: 2043-2051

TierSchNutzV. (2006) Verordnung zum Schutz landwirtschaftlicher Nutztiere und anderer zur Erzeugung tierischer Produkte gehaltener Tiere bei ihrer Haltung, Tierschutz-Nutztierhaltungsverordnung in der Fassung der Bekanntmachung vom 22. August 2006 (BGBl. I S. 2043), die zuletzt durch Artikel 3 Absatz 2 des Gesetzes vom 30. Juni 2017 (BGBl. I S. 2147) geändert worden ist. letzter Aufruf am 23.02.2020 <http://www.gesetze-im-internet.de/tierschnutzv/BJNR275800001.html>

Vanderhasselt R F, Sprenger M, Duchateau L, Tuytens F A M. (2013) Automated assessment of footpad dermatitis in broiler chickens at the slaughter-line: Evaluation and correspondence with human expert scores. Poultry Science, 92: 12-18

Welfare Quality®. (2009) Welfare Quality® Assessment Protocol for Poultry (broilers, Laying

Hens). Lelystad, Netherlands



## X. VERZEICHNIS DER ABBILDUNGEN UND TABELLEN

### 1. Abbildungsverzeichnis

<b>Abbildung 1:</b> Beispielfoto ohne Softwareraster von Füßen, welche in der Schlachtkette durch die festinstallierte Kamera fahren (Bild: CLK GmbH).....	24
<b>Abbildung 2:</b> Beispielfoto mit Softwareraster von Füßen, welche in der Schlachtkette durch die festinstallierte Kamera fahren; grün = Fläche gesamter Fußballen, rot = Fläche der Läsion (Bild: CLK GmbH).....	24
<b>Abbildung 3:</b> Zuordnungsmatrizen des Mikroscores zugeordnet dem Makroscore mit optimistischen / neutralen / pessimistischen Szenario.....	43
<b>Abbildung 4:</b> Schätzung der Power in Bezug auf die Stichprobengröße pro Makroscore-Ausprägung mit optimistischen / neutralen / pessimistischen Szenario. ....	44
<b>Abbildung 5:</b> Kamerascore in Bezug auf die manuell gemessene, relative Flächenveränderung der Foot-Pad Dermatitis. ....	46
<b>Abbildung 6 :</b> Bland-Altman Plot: Die Differenz der relativen Flächenveränderung in % und des Kamerascores in Bezug auf die relative Flächenveränderung der Foot-Pad Dermatitis in %.....	47
<b>Abbildung 7:</b> Verbesselter Kamerascore in Bezug die manuell gemessene, relative Flächenveränderung.....	48
<b>Abbildung 8:</b> Deskriptive Darstellung des Makroscores und Kamerascores im Vergleich (Makroscore 3 zusammengefasst, vorher: Makroscore 3 und 4). ....	49
<b>Abbildung 9:</b> Wahrscheinlichkeitsschätzung für die Verteilung des Makroscores entlang des Kamerascores (n = 200). ....	51
<b>Abbildung 10:</b> Höchste Wahrscheinlichkeiten für die Makroscores in Bezug auf den Kamerascore (Makroscore 3 und 4 als 3 zusammengefasst). ....	52
<b>Abbildung 11:</b> Schätzung bedingten Wahrscheinlichkeiten der histologischen Kategorien (gesund, milde Läsionen oder Ulzerationen) in Abhängigkeit des jeweiligen Kamerascores. ....	55
<b>Abbildung 14:</b> Prognosegüte (Performance) von Kamerascore und Makroscore verglichen mit den histopathologischen Ergebnissen (n = 200, ACC = Genauigkeit, Sens = Sensitivität, Spec = Spezifität, PPV = positiver prädikativer Wert, NPV: Negativer prädikativer Wert,	

value= Wert der Prognosegüte, Performance = Prognosegüte). ..... 56

## 2. Tabellenverzeichnis

<b>Tabelle 1:</b> Makroskopischer Score für Foot-Pad Dermatitis nach Martland (1985)...	12
<b>Tabelle 2:</b> Makroskopischer Score für Foot-Pad Dermatitis nach Martrenchar et al. (2002). .....	13
<b>Tabelle 3:</b> Makroskopischer Score für Foot-Pad Dermatitis nach Ekstrand et al. (1997) verglichen mit makroskopischem Score für Foot-Pad Dermatitis nach Ekstrand und Svedberg (Ekstrand et al., 1997). .....	14
<b>Tabelle 4:</b> Makroskopischer Score für Foot-Pad Dermatitis nach Bilgili et al. (2006). .....	15
<b>Tabelle 5:</b> Makroskopischer Score für Foot-Pad Dermatitis nach Welfare Quality® Assessment protocol for poultry (2009).....	15
<b>Tabelle 6:</b> Mikroskopischer Score für Foot-Pad Dermatitis nach McKeegan (2010).	17
<b>Tabelle 7:</b> Mikroskopischer Score für Foot-Pad Dermatitis nach Michel et al. (2012). .....	18
<b>Tabelle 8:</b> Makroskopischer Score für Foot-Pad Dermatitis nach den Ausführungshinweisen der Tierschutz-Nutztierhaltungsverordnung für Masthühner (2011). .....	19
<b>Tabelle 9:</b> Makroskopischer Score für Foot-Pad Dermatitis nach dem Runderlass des Niedersächsischen Ministeriums für Ernährung, Landwirtschaft und Verbraucherschutz (2014).....	20
<b>Tabelle 10:</b> 5-stufiger Makroscore zur Bewertung von Foot-Pad Dermatitis modifiziert nach dem Welfare Quality® Assessment protocol for poultry (2009) (Bilder: Piller). .....	22
<b>Tabelle 11:</b> Vier-stufiger Microscore zur Bewertung von Foot-Pad Dermatitis modifiziert nach dem Schema von Michel et al. (2012) (Bilder: Piller et al., 2020).....	26
<b>Tabelle 12:</b> Verteilung des Kamerancores entlang der Makroscore-Kategorien (n = 200, Mean= Mittelwert, SD =Standardabweichung, Min = Minimalwert, Max= Maximalwert).....	50
<b>Tabelle 13:</b> Cut-Off-Werte für die Verteilung der Makroscores entlang des Kamerancores (n = 200, Min =Minimalwert, Max = Maximalwert). ....	53
<b>Tabelle 14:</b> Für den Makroscore in Bezug auf die Cut-Off-Werte (Makroscore 3 und 4 zusammengefasst, n = 200, KS = Kamerancore, Acc = Genauigkeit, Sens = Sensitivität, Spec = Spezifität, PPV = positiver prädikativer Wert, NPV: Negativer prädikativer Wert, value= Wert der Prognosegüte, Performance = Prognosegüte).....	54





## XI. ANHANG

**Tabelle I: Erklärungen der statistischen Begriffe der Gütemaße.**

<b>Gütemaß</b>	<b>Erklärung</b>
<b>Sensitivität</b>	Die Wahrscheinlichkeit, dass ein Tier der Kategorie k auch als solches identifiziert wird
<b>Spezifität</b>	Die Wahrscheinlichkeit, dass ein Tier, dass nicht der Klasse k angehört, auch als solches identifiziert wird
<b>Positiver prädiktiver Wert</b>	Die Wahrscheinlichkeit, dass ein Tier, welches der Kategorie k zugeordnet wurde, in Wahrheit auch in diese Kategorie gehört
<b>Negativer prädiktiver Wert</b>	Die Wahrscheinlichkeit, dass ein Tier, welches nicht der Kategorie k zugeordnet wurde, in Wahrheit auch nicht in diese Kategorie gehört
<b>Accuracy/Genauigkeit</b>	Wahrscheinlichkeit für die korrekte Zuordnung der Tiere



## **XII. DANKSAGUNG**

Mein besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr. Dr. Michael Erhard für das Ermöglichen und Überlassen dieser Arbeit, sowie hervorragenden Hilfestellungen und Führung.

Zudem möchte ich mich ganz herzlich bei Frau Dr. Helen Louton für die exzellente Betreuung bedanken. Besonders für die gemeinsame Durchführung sowie die Leitung des Projektes, wissenschaftliche Hilfestellungen, zahlreiches Korrekturlesen, für die konstruktiven Gespräche und Erreichbarkeit auch spätabends und am Wochenende.

Viel Dank geht auch an meine Kolleginnen Frau PD. Dr. Shana Bergmann und Frau Dr. Angela Schwarzer, welche mir immer mit Rat und Tat zur Seite standen, sowohl bei praktischen Fragestellungen als auch bei fachlichen Fragen rund ums Geflügel und den Tierschutz.

An dieser Stelle möchte ich auch noch allen anderen Kollegen und Kolleginnen des Lehrstuhls für Tierschutz, Verhaltenskunde, Tierhygiene und Tierhaltung an der Ludwig Maximilians-Universität München sowie Praktikanten und anderen Doktoranden danken, welche mich bei Ausfahrten oder anderen Belangen unterstützt haben.

Herrn Dr. Paul Schmidt gilt mein Dank für das Erstellen der Grafiken und der hervorragenden statistischen Auswertung.

Herrn PD Dr. Sven Reese, Herrn Dr. Josef Bachmeier, Frau Dr. Verena Lietze, Herrn Dr. Schade, Frau Dr. Böhm, Frau Dr. Kappe und dem Team des Tiergesundheitsdienst Bayern e.V danke ich sehr für die tolle fachliche Unterstützung.

Des Weiteren möchte ich Frau Yeliz und Herrn Stefan Wiesbeck sowie Herrn Matthias Hastreiter vom Schlachthof Bogen für die zuvorkommende Mitarbeit vor Ort am Schlachthof danken

Frau Prof. Dr. Kemper, Frau Dr. Spindler und Frau Dr. Stracke vom Institut für Tierhygiene, Tierschutz und Nutztierethologie der Stiftung Tierärztliche Hochschule Hannover gilt ebenso mein großer Dank sowohl für die Mitarbeit beim Beobachterabgleich, dem Mitwirken an der Publikation sowie auch für den guten fachlichen Austausch.

Der CLK GmbH und im Besondern Herrn Jan Schulte-Landwehr danke ich für die tolle fachliche und technische Unterstützung mit dem Kamerasystem.

Auch den Kolleginnen und Kollegen vom Fachbereich Nutztierethologie und Tierhaltung der Universität Kassel und der Fachgruppe Tierhaltung und Produkte (Studienschwerpunkt

angewandte Geflügelwissenschaften) der Hochschule Osnabrück als Verbundspartner des Autowohlprojekt gilt mein herzlicher Dank für die fachliche Unterstützung und für den tollen wissenschaftlichen Austausch in zahlreichen Treffen und Workshops.

Von ganzen Herzen möchte ich mich bei meiner Familie bedanken: Meinem Bruder Patrick, der mich bedingungslos bereits mein ganzes Leben unterstützt, widme ich diese Arbeit. Ebenso danke ich meiner Mutter Manuela, die sich immer nur das Beste für mich wünscht und meinem Vater Martin ohne dessen großartige Unterstützung ich weder Tierarzt geworden wäre noch diese Doktorarbeit angefertigt hätte. Auch meinen Großeltern, besonders Maria und Johann, welche mir schon als Kind die Liebe zu den Tiere zeigten, gilt mein großer Dank.

Ganz besonders möchte ich noch Simon danken, dessen unermüdliche Unterstützung mir sehr viel Kraft und Motivation während meiner Zeit als Promotionsstudent gegeben hat. Tausend Dank!

Als letztes geht mein Dank an alle Tiere, sowohl an meine „emotional support animals“ den Hund Timmi, den Hamster Tyron und den Meerschweinchen Nacho, Salsa und Tequila sowie auch an die Hühner, deren Füße ich für diese Arbeit untersucht habe. Jedes Tier verdient unseren Respekt und unseren Schutz unabhängig von der Spezies.